VOLUME 106 PART B NUMBER 27 MAY 1959



The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART B

ELECTRONIC AND COMMUNICATION ENGINEERING
(INCLUDING RADIO ENGINEERING)

SAVOY PLACE · LONDON W.C.2

Price Ten Shillings and Sixpence

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN

COUNCIL 1958–1959

President

S. E. GOODALL, M.Sc.(Eng.), F.Q.M.C.

Past-Presidents

W. H. ECCLES, D.Sc., F.R.S.
THE RT. HON. THE EARL OF MOUNT EDGCUMBE, T.D.
J. M. DONALDSON, M.C.
PROFESSOR E. W. MARCHANT, D.Sc.
H. T. YOUNG.
SIR GEORGE LEE, O.B.E., M.C.
SIR ARTHUR P. M. FLEMING, C.B.E., D.Eng., LL.D.
J. R. BEARD, C.B.E., M.Sc.
SIR NOEL ASHBRIDGE, B.Sc.(Eng.).
SIR HARRY RAILING, D.Eng.
P. DUNSHEATH, C.B.E., M.A., D.Sc.(Eng.).
SIR VINCENT Z. DE FERRANTI, M.C.
T. G. N. HALDANE, M.A.
PROFESSOR E. B. MOULLIN, M.A., Sc.D., LL.D.
SIR ARCHIBALD J. GILL, B.Sc.(Eng.).
SIR JOHN HACKING.
COLONEL B. H. LEESON, C.B.E., T.D.
SIR HAROLD BISHOP, C.B.E., B.Sc.(Eng.)
SIR JOSIAH ECCLES, C.B.E., D.Sc.
SIR GEORGE H. NELSON, Bart.
SIR GORDON RADLEY, K.C.B., C.B.E. Ph.D.(Eng.).
T. E. GOLDUP, C.B.E.

Vice-Presidents

SIR WILLIS JACKSON, D.Sc., D.Phil., Dr.Sc.Tech., F.R.S. G. S. C. LUCAS, O.B.E. SIR HAMISH D. MACLAREN, K.B.E., C.B., D.F.C., LL.D., B.Sc.
C. T. MELLING, C.B.E., M.Sc.Tech. A. H. MUMFORD, O.B.E., B.Sc.(Eng.).

Honorary Treasurer

E. LEETE.

Ordinary Members of Council

J. A. BROUGHALL, B.Sc.(Eng.).
PROFESSOR M. W. HUMPHREY DAVIES, M.Sc.
SIR JOHN DEAN, B.Sc.
B. DONKIN, B.A.
J. M. FERGUSON, B.Sc.(Eng.).
D. C. FLACK, B.Sc.(Eng.), Ph.D.
J. S. FORREST, D.Sc., M.A.
R. J. HALSEY, C.M.G., B.Sc.(Eng.).
E. M. HICKIN.
J. B. HIGHAM, Ph.D., B.Sc.
F. C. MCLEAN, C.B.E., B.Sc.
B. L. METCALF, B.Sc.(Eng.).
J. R. MORTLOCK, Ph.D., B.Sc.(Eng.).
R. H. PHILLIPS, T.D.
H. V. PUGH.
J. R. RYLANDS, M.Sc., J.P.
D. P. SAYERS, B.Sc.
C. E. STRONG, O.B.E., B.A., B.A.I.
D. H. TOMPSETT, B.Sc.(Eng.).
H. WATSON-JONES, M.Eng.
H. WEST, M.Sc.

Chairmen and Past-Chairmen of Sections

Measurement and Control:
J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.
*H. S. PETCH, B.Sc.(Eng.).

Electronics and Communications:
G. MILLINGTON, M.A., B.Sc.
*J. S. MCPETRIE, Ph.D., D.Sc.

Supply:
D. P. SAYERS, B.Sc.
*PROFESSOR M. G. SAY, Ph.D., M.Sc., F.R.S.E.

Utilization:
R. A. MARRYAT, B.Sc.(Eng.).
*J. VAUGHAN HARRIES.

Chairmen and Past-Chairmen of Local Centres

East Midland Centre:
D. E. LAMBERT, B.Sc.(Eng.).
*J. D. PIERCE.

Mersey and North Wales Centre:
J. COLLINS.
*T. MAKIN.

North Midland Centre:
J. D. NICHOLSON, B.Sc.
*A. J. COVENEY.

North-Eastern Centre
A. T. CRAWFORD, B.Sc.
*T. W. WILCOX.

Southern Centre:
G. BISHOP, B.Sc.
*L. G. A. SIMS, M.Sc., Ph.D., D.Sc.

North-Western Centre:
PROFESSOR F. C. WILLIAMS, O.B.E., D.Sc., D.Phil., F.R.S.
*F. R. PERRY, M.Sc.Tech.

Northern Ireland Centre:
D. S. MCILLHAGGER, Ph.D., M.Sc.
*C. M. STOUPE, B.Sc.

Western Centre:
R. W. STEEL.
*J. F. WRIGHT.

Scottish Centre:
R. J. RENNIE, B.Sc.
*E. O. TAYLOR, B.Sc., F.R.S.E.

South Midland Centre:
J. ASHMORE.
*L. L. TOLLEY, B.Sc.(Eng.).

* Past Chairman.

MEASUREMENT AND CONTROL SECTION COMMITTEE 1958–1959

Chairman

J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.

Vice-Chairmen

PROFESSOR A. TUSTIN, M.Sc.; C. G. GARTON.

Past-Chairmen

H. S. PETCH, B.Sc.(Eng.); DENIS TAYLOR, M.Sc., Ph.D.

Ordinary Members of Committee

E. W. CONNON, B.Sc.(Eng.), M.Eng.
W. S. ELLIOTT, M.A.
W. C. LISTER, B.Sc.
A. C. LYNCH, M.A., B.Sc.
A. J. MADDOCK, D.Sc.
R. E. MARTIN.
R. S. MEDLOCK, B.Sc.
A. NEMET, Dr.Sc.Techn.
G. A. V. SOWTER, Ph.D., B.Sc.(Eng.).
R. H. TIZARD, B.A.
F. C. WIDDIS, B.Sc.(Eng.).
M. V. WILKES, M.A., Ph.D., F.R.S.

And

The President (*ex officio*).
The Chairman of the Papers Committee.
J. B. HIGHAM, Ph.D., B.Sc. (representing the Council).
W. GRAY (representing the North-Eastern Measurement and Electronics Group).
T. KILBURN, M.A., Ph.D., D.Sc. (representing the North-Western Measurement and Control Group).
G. H. RAYNER, B.A. (nominated by the National Physical Laboratory).
H. M. GALE, B.Sc.(Eng.) (representing the South Midland Electronics and Measurement Group).

ELECTRONICS AND COMMUNICATIONS SECTION COMMITTEE 1958–1959

Chairman

G. MILLINGTON, M.A., B.Sc.

Vice-Chairmen

M. J. L. PULLING, C.B.E., M.A. R. J. HALSEY, C.M.G., B.Sc.(Eng.). T. B. D. TERRONI, B.Sc.

Past-Chairmen

J. S. MCPETRIE, Ph.D., D.Sc. R. C. G. WILLIAMS, Ph.D., B.Sc.(Eng.).

Ordinary Members of Committee

D. A. BARRON, M.Sc.
W. J. BRAY, M.Sc.(Eng.).
J. BROWN, M.A., Ph.D.
PROFESSOR A. L. CULLEN, Ph.D., B.Sc.
C. W. EARP, B.A.
L. I. FARREN, M.B.E.
V. J. FRANCIS, B.Sc.
G. G. GOURIET.
COMDR. C. G. MAYER, O.B.E., U.S.N.R.
J. MOIR.
B. G. PRESSEY, M.Sc.(Eng.), Ph.D.
T. R. SCOTT, D.F.C., B.Sc.
W. E. WILLSHAW, M.B.E., M.Sc.Tech.
A. J. YOUNG, B.Sc.(Eng.).

And

The President (*ex officio*).
The Chairman of the Papers Committee.
C. E. STRONG, O.B.E., B.A., B.A.I. (representing the Council).
E. H. COOKE-YARBOROUGH, M.A. (Co-opted Member).
G. H. HICKLING, B.Sc. (representing the North-Eastern Measurement and Electronics Group).
N. C. ROLFE, B.Sc.(Eng.) (representing the Cambridge Electronics and Communications Group).
K. J. BUTLER (representing the North-Western Electronics and Communications Group).
R. E. YOUNG, B.Sc.(Eng.) (representing the South Midland Electronics and Measurement Group).

The following nominees:
Royal Navy: CAPTAIN R. L. CLODE, R.N.
Army: COL. R. G. MILLER, M.A.
Royal Air Force: GROUP CAPTAIN A. FODEN, B.Sc.Tech., R.A.F.

Secretary

W. K. BRASHER, C.B.E., M.A., M.I.E.E.

Deputy Secretary

F. JERVIS SMITH, M.I.E.E.

Principal Assistant Secretary

F. C. HARRIS.

Editor-in-Chief

G. E. WILLIAMS, B.Sc.(Eng.), M.I.E.E.

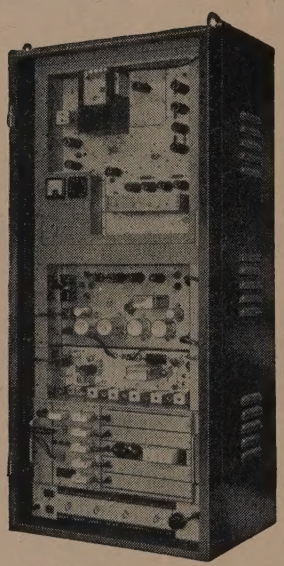
Introducing the MC10

a NEW 5-Channel Transistorised VHF Radio-Telephone Terminal



A TYPICAL LOCATION where this new 5-Channel Transistorised VHF Radio-Telephone Terminal will prove indispensable in vital communications

Compact design which at low cost provides five high-grade telephone circuits



Radio

Frequency range 156 - 184 Mc/s
(other frequency ranges are available)
Transmitter power output 30 watts
Deviation 75 kc/s
Receiver Noise Factor 8db
All characteristics of the transmitter and receiver conform to CCIR specifications

Carrier Telephone Equipment

5 Telephone channels · 4 kc/s spaced
Equipped with Out of Band Signalling
Facilities for dialling, Ringdown or junction working
Printed Wiring · Plug-in Units
Crystal frequency control
Resin cast components

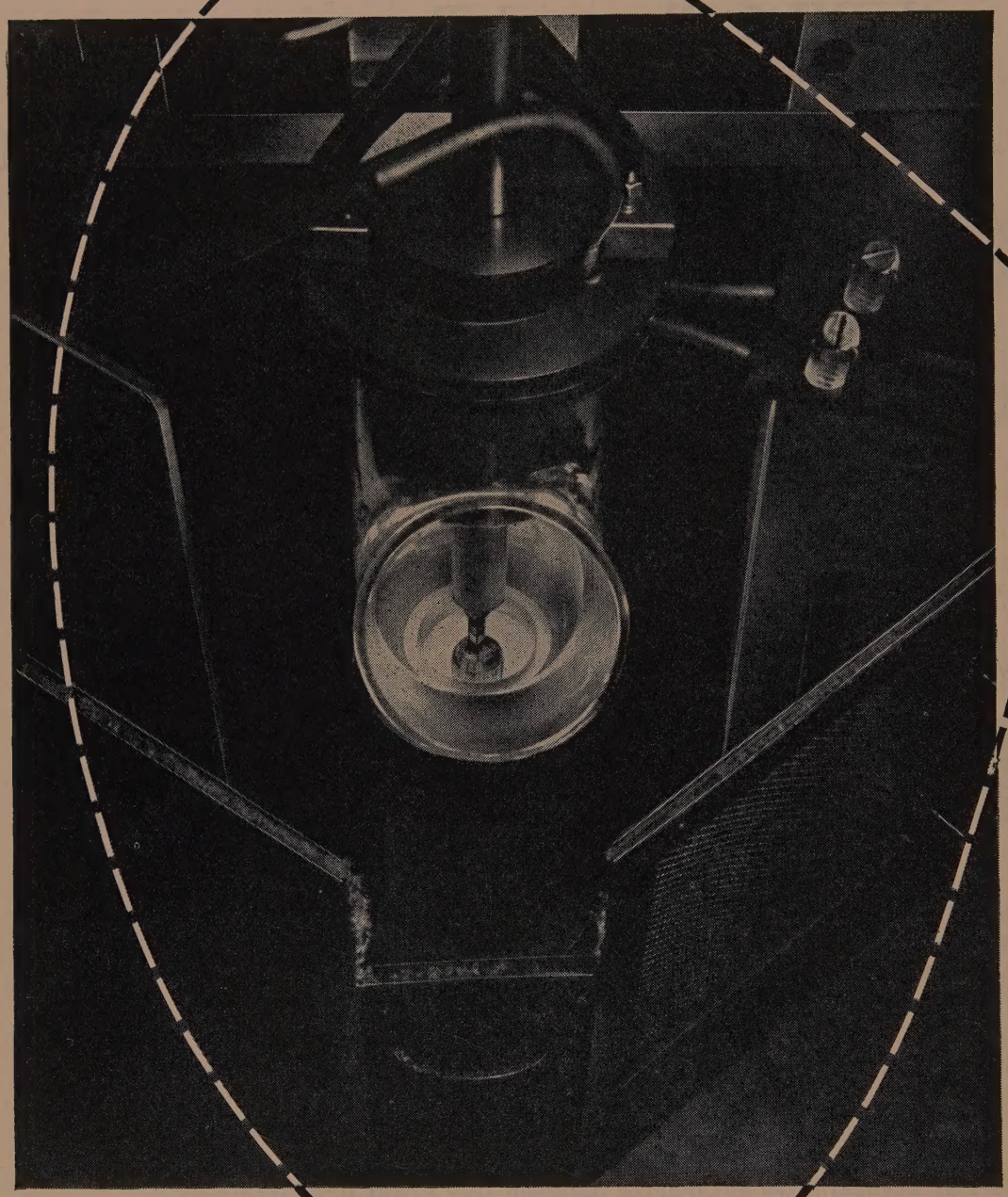
The result of co-operative enterprise between two great organizations

Rediffon — SIEMENS EDISWAN

REDIFON LIMITED
Communications Division, Wandsworth, London, S.W.18
Telephone : VANdyke 7281
A Manufacturing Company in the Rediffusion Group

SIEMENS EDISON SWAN LIMITED
Transmission Division, Woolwich, London, S.E.18
Telephone : Woolwich 2020
An A.E.I. Company

The heart of the matter...



The formation of a single silicon crystal ingot

FERRANTI

FERRANTI LTD · GEM MILL · CHADDERTON · OLDHAM · LANCS · *Telephone:* MAIn 6661

London Office: KERN HOUSE · 36 KINGSWAY · W.C.2 *Telephone:* TEMple Bar 6666

...the art of the matter

FERRANTI

offer the widest range of

SILICON Semiconductor Devices

in the United Kingdom

Ferranti Ltd. were the first company in Britain to introduce Silicon semiconductor devices as used in magnetic amplifiers, in aircraft, guided missiles, radar and computers. Until recently they were the only firm in the United Kingdom supplying silicon diodes in quantity. Commencing their programme of research and development in 1954, they have already made outstanding contributions to technique, and are now producing at Gem Mill, Oldham well over half-a-million silicon diodes annually in the widest range offered by any British manufacturer.

Among the 120 or more different devices are rectifiers, fast diodes, zener diodes, tetralayer diodes and triodes, alloy junction transistors, diffused junction transistors, photovoltaic cells, voltage variable capacitors and many new ones.

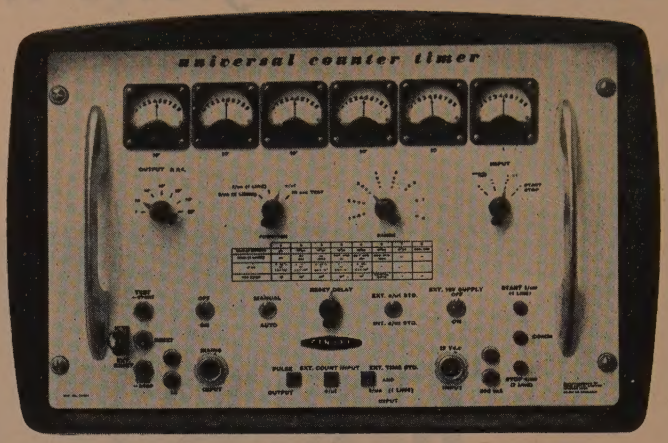
Data Sheets, Application Reports etc., advice and assistance in techniques of application are freely available.



Visit the **International Transistor Exhibition** promoted by the Electronics and Communications Section of The Institution of Electrical Engineers, 21-27 May, at Earls Court, London. Admission by ticket only, available from Ferranti Ltd. or from the I.E.E.

FE191

Transistorized UNIVERSAL COUNTER TIMER



Frequency Measurement

Random Counting

Frequency Division

Time Measurement

Frequency Standard

This fully transistorized portable equipment provides for a wide range of time and frequency measurement as well as facilities for counting, frequency division and the provision of standard frequencies. The facilities available are briefly listed below:

TIME/UNIT EVENT (1 LINE): For the measurement of the time interval between two occurrences in a continuously varying electrical function in the range 3μsec to 1sec. The time for 1, 10 or 100 such events can be measured.

TIME/UNIT EVENT (2 LINE): For time measurement in range 1μsec to 2777hrs. of any interval defined by a positive or negative going pulse in any combination.

EVENTS/UNIT TIME: For frequency measurement in range 30c/s to 1Mc/s over period of 0·001, 0·01, 0·1, 1 or 10secs. Crystal accuracy ±2 parts in 10^6/week. For mains or 12Vd.c. operation.

Full technical specification available on request.

CINTEL

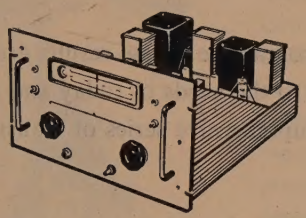
RANK CINTEL LIMITED
WORSLEY BRIDGE ROAD · LONDON · SE26
HITHER GREEN 4600

Sales and Servicing Agents: Atkins, Robertson & Whiteford Ltd. Industrial Estate, Thornliebank, Glasgow; McKellen Automation Ltd., 122 Seymour Grove, Old Trafford, Manchester, 16; Hawnt & Co. Ltd., 59 Moor St., Birmingham, 4.

A.W.A ELECTRONICS

New concepts in electronics have been developed at AWA, as a result of experience with missile systems. Now they have a wider application. Here are some of the new AWA devices now available to industry.

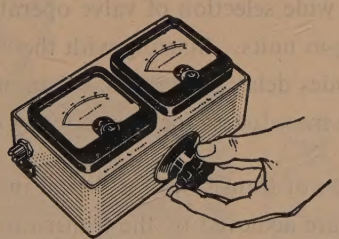
U.H.F. RECEIVER



Designed as a Wide Band Low Noise Receiver for the 420/470 Mc/s Frequency Band. Available as either a 19″ rack mounted unit c/w Stabilised Power Unit, or in an 8″ Case with separate Power Unit. Basic arrangement consists of R.F. Amplifier, Mixer, Local Oscillator I.F. Amplifier (A.G.C. Controlled) Cathode Follower Output Stage. Wide variations of this Receiver can be supplied to customers' own requirements.

Standard Specification: Frequency Range: 420/470 Mc/s; Bandwidth: 4.5 Mc/s; Noise Factor: 10db (approx.); Intermediate Frequency: 45 Mc/s; Sensitivity: 25 MicroV. for a 12 db signal to noise ratio; R.F. Gain: 12 db; I.F. Gain: 80 db; Image Rejection: 40 db; Input Impedance: 75 ohms (approx.) Unbalanced; Output Impedance: 80 ohms (approx.); Outputs: (a) 0.5v from Crystal Detector, (b) 300 mV at 45 Mc/s.

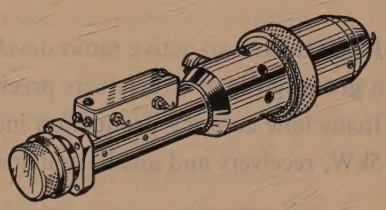
DIRECTIONAL COUPLER



Of the "Loop" type, suitable for measurements of R.F. power and Standing Wave Ratio in coaxial cables. Directional properties are largely unaffected by frequency changes, so coupler may be used to help obtain optimum termination for a 52 ohm coaxial system up to 600 Mc/s.

72 ohm version of this instrument is also available.

Standard Specification: Case Size: 7″ x 4″ x 2½″; Weight: 4 lbs. 3 ozs.; Power Measurement: low range 1w cw max., high range 5w cw max.; Accuracy (at frequency of calibration): low range 0.1 dbs, high range 0.2 dbs; Directivity: 26 dbs (approx.); Coupling Coefficient: 30 dbs (approx.).



PRECISION OSCILLATOR

The Oscillator has been designed round a disc sealed triode and particular attention has been paid to ensure good frequency stability. This has been achieved by the use of selected materials and concentric sleeve tuning. The latter in conjunction with Micrometer head tuning giving very good resolution. In order to reduce R.F. losses all cavities and lines are silver plated, polished, and rhodium flashed.

The Oscillator can be supplied in the form illustrated for installation in the customers' own equipment, or as a unit complete with its own stabilised power supply mounted on a 19″ panel.

Standard Specification: Frequency: Adjustable to operate in the 450/550 Mc/s band. Actual Tuning Range in this band is approx. 30 Mc/s; Frequency Stability: Better than 1 part in 10^4 (long term); Input: 300v at 30 mA, 6.3v at 0.4A; Power Output: Max. output 1.25w at 470 Mc/s.

All devices are adaptable to suit customers' own requirements. For further information consult:

COMMERCIAL ELECTRONICS DEPT.

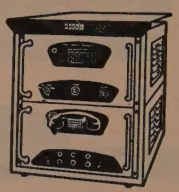
SIR W. G. ARMSTRONG WHITWORTH AIRCRAFT LTD.,
Bagington, Coventry, England.

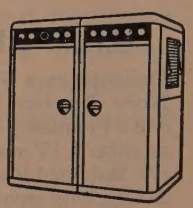
MEMBER OF THE HAWKER SIDDELEY GROUP

ENGINEERING by RACAL

RADIO COMMUNICATIONS







RACAL ENGINEERING LTD., one of the more recent arrivals in the field of radio communications, have already made a name for original design and high quality. Their RA. 17 HF Communications Receiver has established a world-wide reputation and is being adopted in increasing numbers by the top communications authorities.

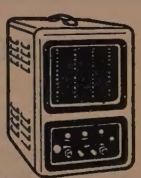
RACAL are ready to supply the increasing demand for Single Sideband radio links needed in modern communication systems. The TRA. 55 Transmitter/Receiver makes available even to small scale users all the advantages of SSB with remarkable simplicity of operation.

RACAL have an active radio development programme whose results are to be seen in a growing range of products precisely suited to the needs of today and tomorrow. Items now entering production include a comprehensive series of transmitters up to 5kW, receivers and ancillary items.

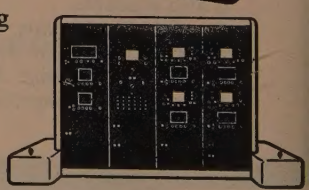
Write to

RACAL ENGINEERING LIMITED

DIGITAL INSTRUMENTATION







RACAL INSTRUMENTS LIMITED offer a wide selection of valve operated Digital Counting Instruments assembled from plug-in units. Starting with the portable SA. 20 equipment, the range includes delay generators, chronometers and frequency meters supplemented by suitable transducers for the majority of applications.

RACAL are well to the fore in the utilization of transistors for digital instrumentation. New standards of reliability and flexibility are achieved by the transistor-ferrite core plug-in cards employed in the construction of their latest range of instruments.

RACAL specialise in the design and manufacture of instrumentation systems employing standard equipment on a building-block basis. Their engineers, who are always available to discuss the problems and requirements of the factory and the laboratory, will be glad to provide unbiased advice and submit proposals.

Write to

RACAL INSTRUMENTS LIMITED

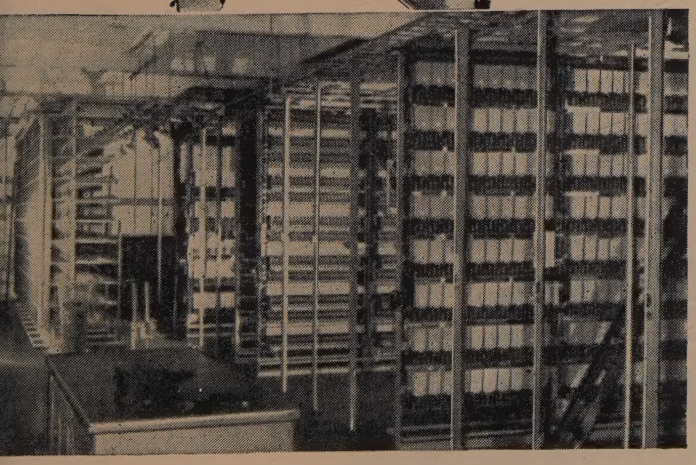
WESTERN ROAD, BRACKNELL, BERKS Telephone : Bracknell 941 Telegrams/Cables : RACAL BRACKNELL BERKS
N. ENGLAND AGENT : Farnell Instruments, Ltd. Wetherby Industrial Estate, York Road, Wetherby, Yorks. Tel : Wetherby 2691/2
SCOTTISH AGENT : A. R. Bolton & Co., Ltd., 3A St. Vincent Street, Edinburgh. Tel : Edinburgh 32035
OVERSEAS : Agents operate in the majority of territories throughout the world



MODERN COMMUNICATIONS IN THE ARGENTINE

The Argentine Republic has taken a further big step in modernising its communication services. Posadas, a city in the extreme north east of the Republic, has been equipped with a new 3,600 line telephone exchange. A.T.E. have supplied this exchange through their associate T.A.P.E.S.A., as a part of a comprehensive plan to augment the local service and extend the long distance circuits in the area. It was contracted to the requirements of the state Telephone Undertaking (Empresa Nacional de Telecommunicaciones). Its installation is yet another example of the advanced techniques in telecommunication which A.T.E. are supplying to countries throughout the world.





A view of the automatic equipment in the new A.T.E. exchange. A.T.E. supplied the first automatic exchange in Argentina in 1913 at Cordoba.

AUTOMATIC TELEPHONE & ELECTRIC CO. LIMITED

Strowger House, Arundel Street, London, W.C.2.

Distributors in the Argentine: TELEFONOS AUTOMATICOS & PRODUCTOS ELECTRICOS S.A, San-Martin. 574-6° PISO. Buenos Aires.

AT7141

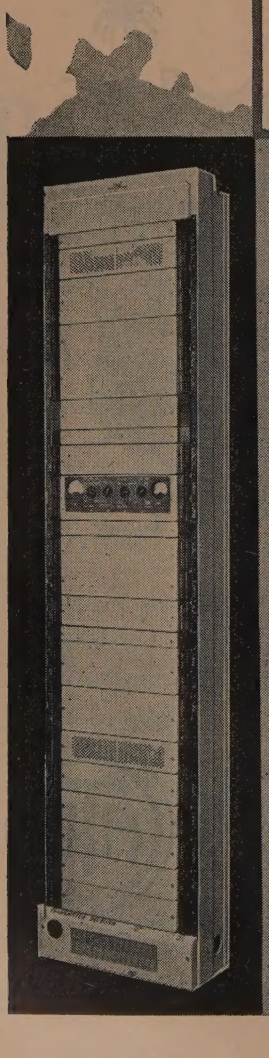
G.E.C. Micro

FIRST IN NIGERI



NOW

The General Electric Com
is to supply a m
circuit microwave radio sy
between Saint John
Moncton, for The New Bruns
Telephone Company, Lim
G.E.C. 240-circuit 2000
radio equipment w
employed, equipped ini
for 60 speech cir
The equipment will co
television signals inste
telephony circuits, if requ
Three intermediate sta
will be provided at Clare



240 circuit
terminal rack

G.E.C.

Everything for Telecommunications

THE GENERAL ELECTRIC COMPANY LIMITED

TELEPHONE, RADIO AND TELEVISION WORKS · COVENTRY · ENGLAN

ve Links

IRST IN LIBYA

CANADA

n Mountain and
t Boyd. A standby link
e supplied with
atic changeover in the
of failure or
lation of the main signal.
er 240-circuit systems
added as required
the same standby link
rotection channel.
formation on the radio
ultiplexing equipment,
write for Standard
cations SPO.5501
PO.1370.

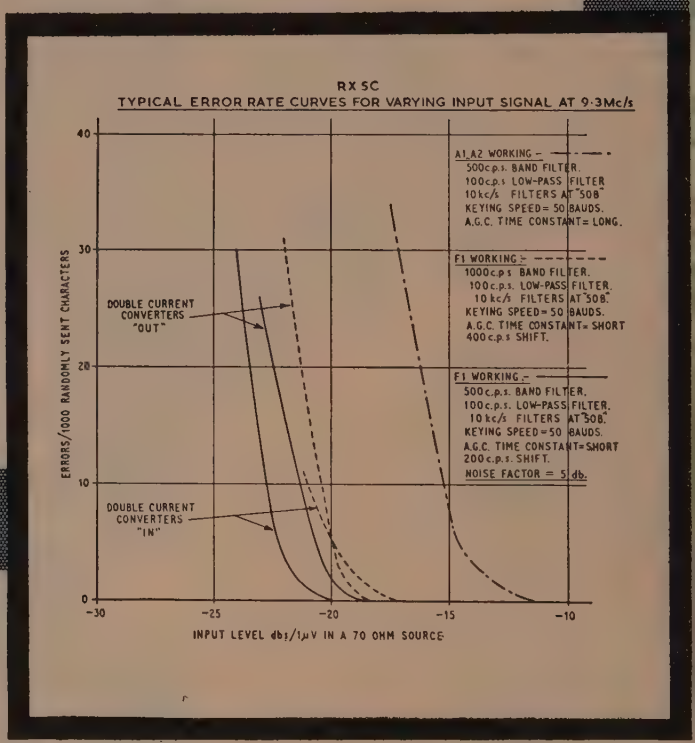


ND

GEC 23

Conventional F.S.K. receivers outmoded by new Standard RX5C

- *The Standard RX5C represents probably the most notable advance in F.S.K. receiver design in the last decade.*
- *The whole concept centres round greatly increased "good-printing" time when compared to conventional limiter/discriminator techniques.*
- *Intensive study of field conditions has enabled some welcome technical features to be incorporated of particular interest to the discerning operator.*
- *The equipment is fully tropicalised to current governing user-specifications.*
- *A range of passband filters from 0.5 kc/s to 6 kc/s, variable frequency oscillator and Four-Frequency Diplex operation (F.6) are optionally available.*
- *Special attention has been paid to the use of "common valve" types and full self-checking circuit arrangements are incorporated to enable the operator to align and set up the receiver without external test apparatus.*



SPECIAL SYSTEMS GROUP

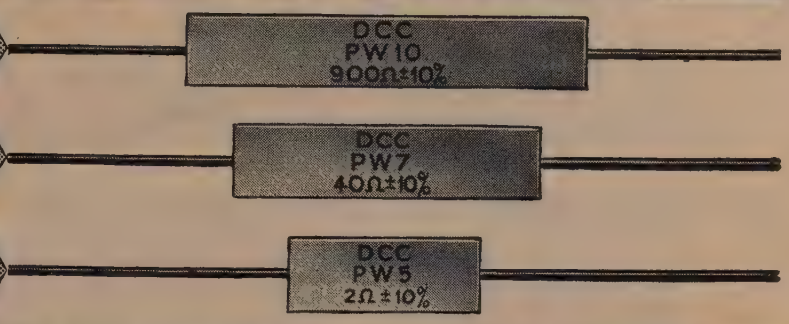
Standard Telephones and Cables Limited

Registered Office: Connaught House, Aldwych, London, W.C.2

RADIO DIVISION: OAKLEIGH ROAD · NEW SOUTHGATE · LONDON N.11

FOR THE FIRST TIME IN BRITAIN

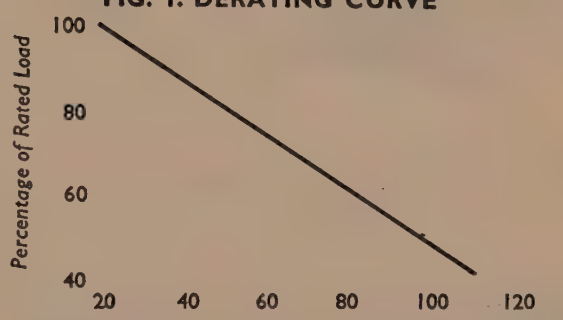
New POWER WIRE-WOUND RESISTORS!



HIGH GRADE RESISTORS AT LOW COST

A new Dubilier process makes available to the design engineer a power wire-wound resistor possessing high-grade characteristics which costs no more than an equivalent standard type. The resistance wire is uniformly wound on a silicone-processed fibre-glass core which is then sealed into a ceramic housing. The result is a remarkably stable resistor which is completely insulated except for the connecting wires.

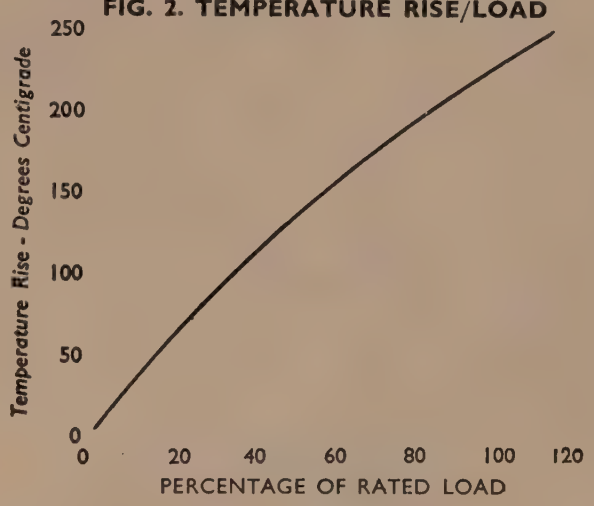
FIG. 1. DERATING CURVE



PERFORMANCE UNDER OPERATING CONDITIONS

- ✳ Resistance change less than 5% after 100 hours at 40°C. ambient temperature and 95% relative humidity.
- ✳ Resistance change less than 2% after three times normal load for 5 seconds.
- ✳ Resistance change less than 5% after 500 hours at full load in 25°C. ambient temperature.
- ✳ Resistance change less than 1% and no physical effects due to soldering.

FIG. 2. TEMPERATURE RISE/LOAD



MAXIMUM TEMPERATURE COEFFICIENT BETWEEN – 55 AND +275°C.

TYPE	0.05%/°C.	0.03%/°C.
PW5	0.5Ω to 2.5Ω	2.5Ω to 2.0kΩ
PW7	0.5Ω to 8.0Ω	8.0Ω to 6.5kΩ
PW10	1.0Ω to 10Ω	10Ω to 10kΩ

TYPE	PW5	PW7	PW10
Wattage	5.0	7.0	10.0
Min. Value	0.5Ω	0.5Ω	1.0Ω
Max. Value	2.0kΩ	6.5kΩ	10kΩ
Length	$\frac{7}{8}$"	$1\frac{25}{64}$"	$1\frac{7}{8}$"

Width and height of all three types are $\frac{3}{8}$" and $\frac{11}{32}$" respectively.

DUBILIER

Catalogue R15A available on request.

DUBILIER CONDENSER CO. (1925) LTD · DUCON WORKS · VICTORIA ROAD · NORTH ACTON · LONDON W.3.

TELEPHONE: ACORN 2241

TELEGRAMS: HIVOLTCON WESPHONE LONDON

DN 235

Semiconductors Limited

Applications Laboratory

making new equipment possible—existing equipment better . . .



The Semiconductors Limited Transistor Applications Laboratory, with greatly improved facilities, is now installed in a new building adjacent to the Swindon factory. This will considerably increase its capacity to assist manufacturers in the correct selection and use of Semiconductors transistors. You are invited to take advantage of this service, either directly or through your Semiconductors representative.

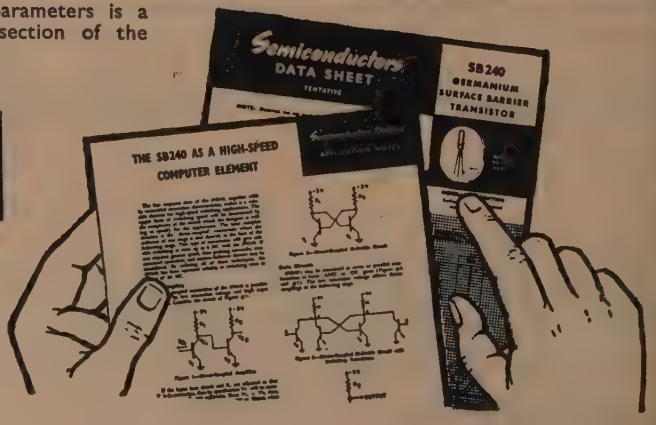
In addition to providing direct assistance to manufacturers, the Applications Laboratory is continuously investigating new approaches to existing problems.

For immediate assistance on your transistor problems telephone the Applications Laboratory SWINDON 6421.

transistors



The evaluation of transistor high-frequency parameters is a continuing study in this specially-equipped section of the Transistor Applications Laboratory.



Semiconductors Limited

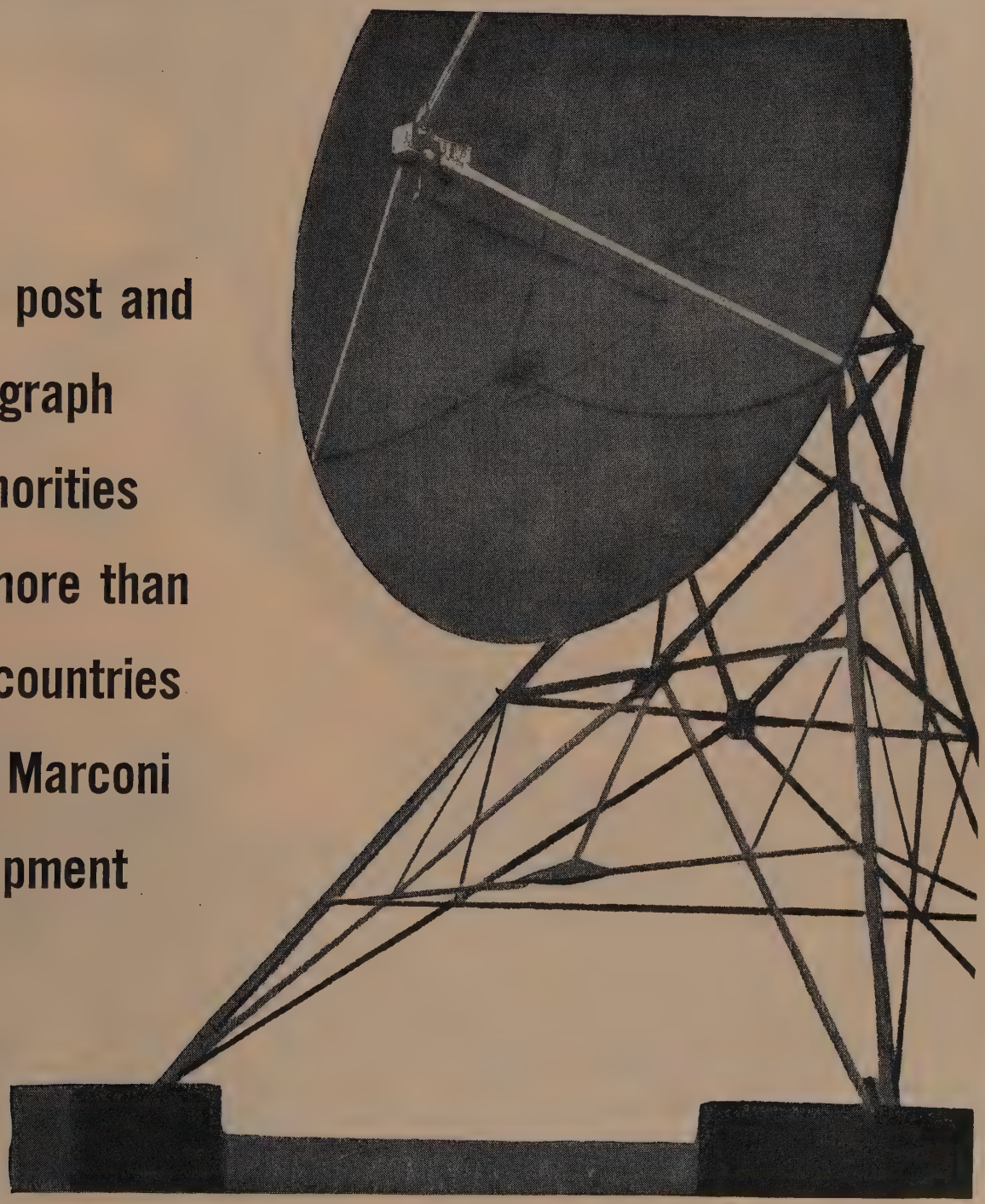
CHENEY MANOR · SWINDON · WILTSHIRE

Telephone: Swindon 6421/4 Telegrams: Semicon, Swindon

** NEW APPLICATION NOTES AND DATA SHEETS ARE CONSTANTLY BEING ISSUED*
ARE YOU ON THE MAILING LIST?

SC.12

Marconi in Telecommunications



The post and telegraph authorities of more than 80 countries use Marconi equipment

MARCONI

COMPLETE COMMUNICATION SYSTEMS

SURVEYED · PLANNED · INSTALLED · MAINTAINED

MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED, CHELMSFORD, ESSEX, ENGLAND

M.1


S.T.C. WORLD-W
S.T.C. are supplying main line microwave telephone systems to 16 countries and have already supplied systems with a capacity of over one million telephone circuit miles, 2000 television channel miles and an equal capacity of standby equipment.
CORNER BROOK
NEWFOU
70-mile over-water path employing the latest automatic diversity system in which the SHF signals are combined in the correct phase before entry into the receiver.
TERMINAL
REPEATER
600-CIRCUIT CAPACITY 2-WAY TE
1-WAY TELEVISION CHANNEL
2-WAY TELEPHONE & TELEVISION
RED ROCKS
CABOT
STRAIT
CAPE NORTH
NOVA SCOTIA
SYDNEY
0 20 40 60 80 100 MILES
0 20 40 60 80 100 120 140 160 KILO
One of 1
S.T.C. MIC
STC
ELECTRONIC SYSTEMS GROUP


E EXPERIENCE
LAND
are supplying a 4000 Mc/s multi-channel microwave
e and television system to the Canadian National
s. It will form the final link in the trans-Canadian
n network. The 536-mile route between St. John's
ney, Nova Scotia, includes a 70-mile over-water path,
ng dual diversity reception, across the Cabot Strait.
ard Telecommunications Laboratories, England, in
ion with S.T.C. (Canada) carried out the propaga-
s on the route and S.T.C. (London) in collaboration
eir Canadian associates will provide technical
e during installation.
are also supplying:
7000 Mc/s Microwave Links to Television Studios
Supervisory and Remote Control Equipment
Antennae
Test Equipment
Entrance Cables
GANDER
ATLANTIC OCEAN
CLARENVILLE
ST. JOHN'S
ountries having
WAVE SYSTEMS
Standard Telephones and Cables Limited
Registered Office Connaught House, Aldwych, London, W.C.2
TRANSMISSION DIVISION: NORTH WOOLWICH · LONDON E.16



in Malaya



The Telecommunications Department of the Government of the Federation of Malaya and Singapore has selected our **1 VF** Signalling and Dialling System for operator dialling on routes radiating from main trunk centres in its rapidly expanding telephone network. This equipment forms an important part of the telecommunications system in Malaya which is recognised as the most up-to-date in this part of the world.

Our **1 VF** system will ensure speedy and reliable trunk connections on these routes at a very low initial cost.

SIEMENS EDISON SWAN LTD
An A.E.I. Company
Telecommunications Division P.D.5
Woolwich, London S.E.18
Telephone: Woolwich 2020

SIEMENS EDISWAN

The best 75 volt stabilisers in the world



BRITISH SERVICES PREFERRED TYPE

M8225/CV4080

The high performance of the Mullard stabiliser 75C1 has led to the recent adoption of its Special Quality equivalent M8225/CV4080 by the British Services as their Preferred 75-volt stabiliser. The M8225/CV4080 is tested for specialised applications in which conditions of extreme shock and vibration are encountered.

- **Wide Current Range . . .**
 2 to 60 milliamps
- **Small Regulation Voltage . . .**
 Less than 9 volts
- **High Stability . . .**
 Typical variation in burning voltage less than ± 2% in any 10,000 hours of operation.



GENERAL PURPOSE TYPE

75C1

The 75C1 is the best 75 volt stabiliser available in the world for general purpose use in industry and communications. It has the same electrical characteristics as the M8225/CV4080 and like this British Services Preferred valve provides an exceptional *combination* of long life, stability and good regulation.

Full data is readily available from the address below.

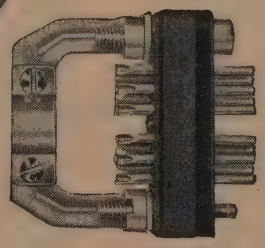
MULLARD LIMITED
MULLARD HOUSE
TORRINGTON PLACE
LONDON · W.C.1
TEL: LANGHAM 6633

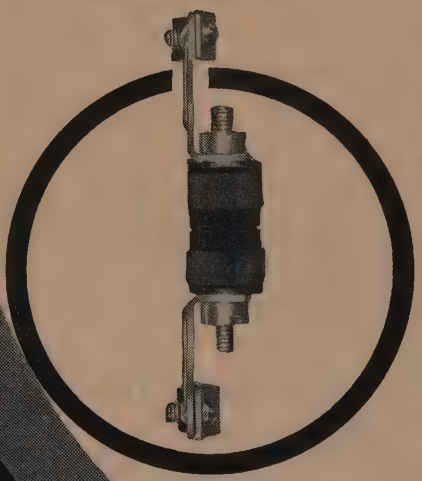
Mullard
GOVERNMENT AND INDUSTRIAL VALVE DIVISION

See Mullard valves & tubes on Stand 2 at the ELECTRONIC COMPUTER EXHIBITION · Olympia November 28 to December 4

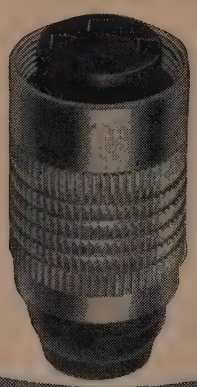
MVT355

a firm GRIP on complex connection









From their wide selection of plugs and sockets, MODAC offer a standard range of rectangular and circular miniature connectors, to which is soon to be added a sub-miniature range with combinations of 10, 16, 28 and 34 contacts.

The main feature of this advanced design is the independently sprung, multi-point contact which provides self-cleaning, low contact resistance and is ideal under conditions of vibration.

MODAC is equipped for the design of connectors to customers' specifications and to supply any of their range as cable assemblies.

They are more than willing—at any time—to come to grips with your general wiring problems.

Modac miniature and sub-miniature connectors

'Modac' is a Registered Trade Mark

for further information and a copy of the 'Modac' catalogue, please contact:

Modern Acoustics Limited
Contact Connectors Division

MANOR WAY • BOREHAM WOOD • HERTS • ELSTREE 3636

Agents for British Commonwealth Countries: Plessey International Limited • Ilford • Essex • Ilford 3040





MAL 2



British-made Newmarket Transistors hold the answer



With the advent of transistors, the solution of ballistic and space travel problems has been enormously speeded up and simplified. Their further development in the next decade may well ordain the shape of many projects yet to come.

If past achievement be any guide, the source from which most future advances will come is predictable. Throughout the *15 years of semiconductor evolution in Britain to date, one name stands out. Besides being the first British company to make available a full range of transistors in production quantities, Newmarket Transistors Ltd were: FIRST with RF types; FIRST again with high power output types; FIRST with intermediate power types; FIRST to supply radio manufacturers with complete matched complements of transistors. In addition, early liaison with leading radio and equipment manufacturers enabled Newmarket to supply the transistors used in the first all-transistor portable radio, the first transistorized loud-hailer, the first transistorized car radio.

Thus, and in every possible way, have Newmarket transistors been tested, proved and improved. The calibre of Newmarket's development background, production resources and customer service is without equal in the industry.

When buying transistors, the name Newmarket is your guarantee of dependability—both in performance and supply.

* *A Bibliography published by the Company is available on application.*

INTERNATIONAL TRANSISTOR EXHIBITION & CONVENTION
Earls Court May 21st—27th 1959

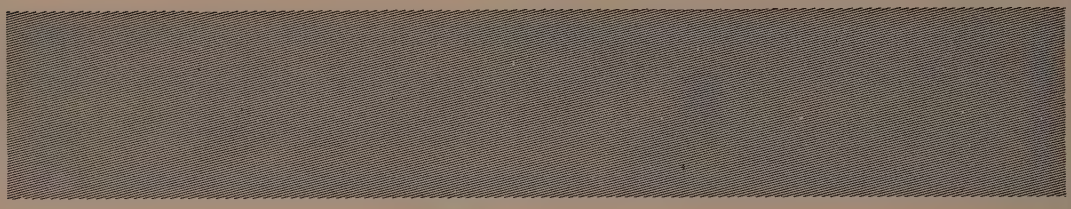
Anyone interested in new applications and developments of transistors should not fail to visit our stand at this Exhibition.

Newmarket Transistors Limited
Exning Road, Newmarket, Suffolk Telephone Newmarket 3381

NEWMARKET TRANSISTORS

TA1640

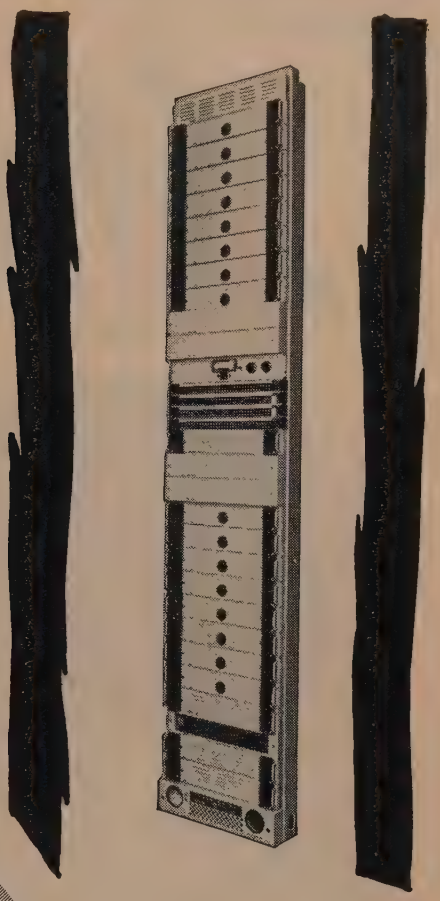
Two new TMC carrier telephone equipments



for S.T.D. and T.A.T.

TELEPHONE MANUFACTURING COMPANY LIMITED

Transmission Division, Cray Works, Sevenoaks Way, Orpington, Kent



FOR LOW-COST BEARERS—
8 CIRCUITS IN 60-108 KC/S BAND
WITH OUT-BAND SIGNALLING

- 6 kc/s channel spacing
- Channel bandwidth 300 c/s—3,400 c/s
- Dialling (E & M) or ring-down facilities
- Low-level signalling (-20dbm0) tone-on or tone-off idle
- Five 8-circuit groups on one double-sided 9 foot rack including carrier, signalling and power supplies
- Alternative version for 4-wire balanced-pair cables with complete equipment for two 16-circuit terminals on one double-sided 9 foot rack

Similar to equipment used by
THE BRITISH POST OFFICE
to provide **SHORT TRUNK CIRCUITS**
with **OUT-BAND DIALLING**



FOR HIGH-COST BEARERS—
16 CIRCUITS IN 60-108 KC/S BAND
WITH HIGH GRADE CHANNELS

- 3 kc/s channel spacing
- Channel bandwidth 300 c/s—3,100 c/s
- Two 16-circuit groups on one double-sided 9 foot rack

Developed by
THE BRITISH POST OFFICE
for use on the
TRANS-ATLANTIC TELEPHONE CABLE

THESE NEW CARRIER TELEPHONE EQUIPMENTS ARE THE LATEST ADDITIONS TO THE T.M.C. RANGE OF TRANSMISSION EQUIPMENT WHICH INCLUDES

Carrier Telephone and Telegraph Equipment for Cable, Open-Wire and Radio
Transistorised 120 c/s and 170 c/s F.M. Telegraph Equipment
Transistorised V.F. Repeater Equipment
Transistorised Privacy Equipment

SELLING AGENTS

AUSTRALIA AND NEW ZEALAND — **Telephone Manufacturing Co. (A'sia) Pty. Ltd., Sydney, N.S.W.**

CANADA AND U.S.A. — **Telephone Manufacturing Co. Ltd., Toronto, Ont.**

ALL OTHER COUNTRIES *(for Transmission Equipment only)* — **Automatic Telephone and Electric Co. Ltd., London**

Marconi

VACUUM-TUBE VOLTMETER TF 1300



Combines Performance & Economy

The TF 1300 is designed to meet the need for a reliable voltmeter of medium range, good accuracy and good stability—at a moderate price. It measures a.c. up to 100 volts in the frequency range 20 c/s to 300 Mc/s, d.c. up to 300 volts, and resistances from 50 ohms to 5 MΩ. The indicating meter is direct-reading on all ranges, for all measurements; no correction factors are necessary. Zero stability is of a high order, and only one zero setting is required for all a.c. or all d.c. ranges.

Embodying a radically new concept of mechanical simplicity which not only reduces production costs but also facilitates servicing, the TF 1300 is an outstanding achievement. For further details, please write for leaflet K141.

£47—PROMPT DELIVERY

AM & FM SIGNAL GENERATORS • AUDIO & VIDEO OSCILLATORS
FREQUENCY METERS • VOLTMETERS • POWER METERS
DISTORTION METERS • FIELD STRENGTH METERS
TRANSMISSION MONITORS • DEVIATION METERS
OSCILLOSCOPES, SPECTRUM & RESPONSE ANALYSERS
Q METERS & BRIDGES

Please address enquiries to MARCONI INSTRUMENTS LTD. at your nearest office:

London and the South: Marconi House, Strand, London, W.C.2 Telephone: COVent Garden 1234

Midlands: Marconi House, 24 The Parade, Leamington Spa Telephone: 1408

North: 23/25 Station Square, Harrogate Telephone: 67455

Export Department: Marconi Instruments Ltd., St. Albans, Herts. Telephone: St. Albans 56161

TC141

'ENGLISH ELECTRIC'

'C' CORES



for smaller, lighter transformers and chokes 'ENGLISH ELECTRIC' 'C' type transformer cores are made from high-quality cold-rolled grain oriented silicon steel strip. A wide range of cores is available and these have been developed to meet the exacting standards of the electronic industry, for smaller, lighter and more efficient transformers and chokes.

Some advantages of using ENGLISH ELECTRIC *'C' cores:—*

Easy to store and handle, their use results in reduced assembly time.
Transformers are up to 30% smaller and lighter.
Transformers have lower losses, lower magnetising VA and lower leakage flux.
Cores are suitable for oil-filled, open-type or resin-cast transformers.
Cores are stable up to 250°C.
Cores are available from stock.

full details from

The English Electric Company Limited, Transformer Sales & Contracts Dept.,
East Lancashire Road, Liverpool, 10. Telephone: AINtree 3641

HE ENGLISH ELECTRIC COMPANY LIMITED, MARCONI HOUSE, STRAND, LONDON, W.C.2.
WORKS: STAFFORD · PRESTON · RUGBY · BRADFORD · LIVERPOOL · ACCRINGTON
.39



TRANS-TAGUS

MICROWAVE TELEPHONE LINK

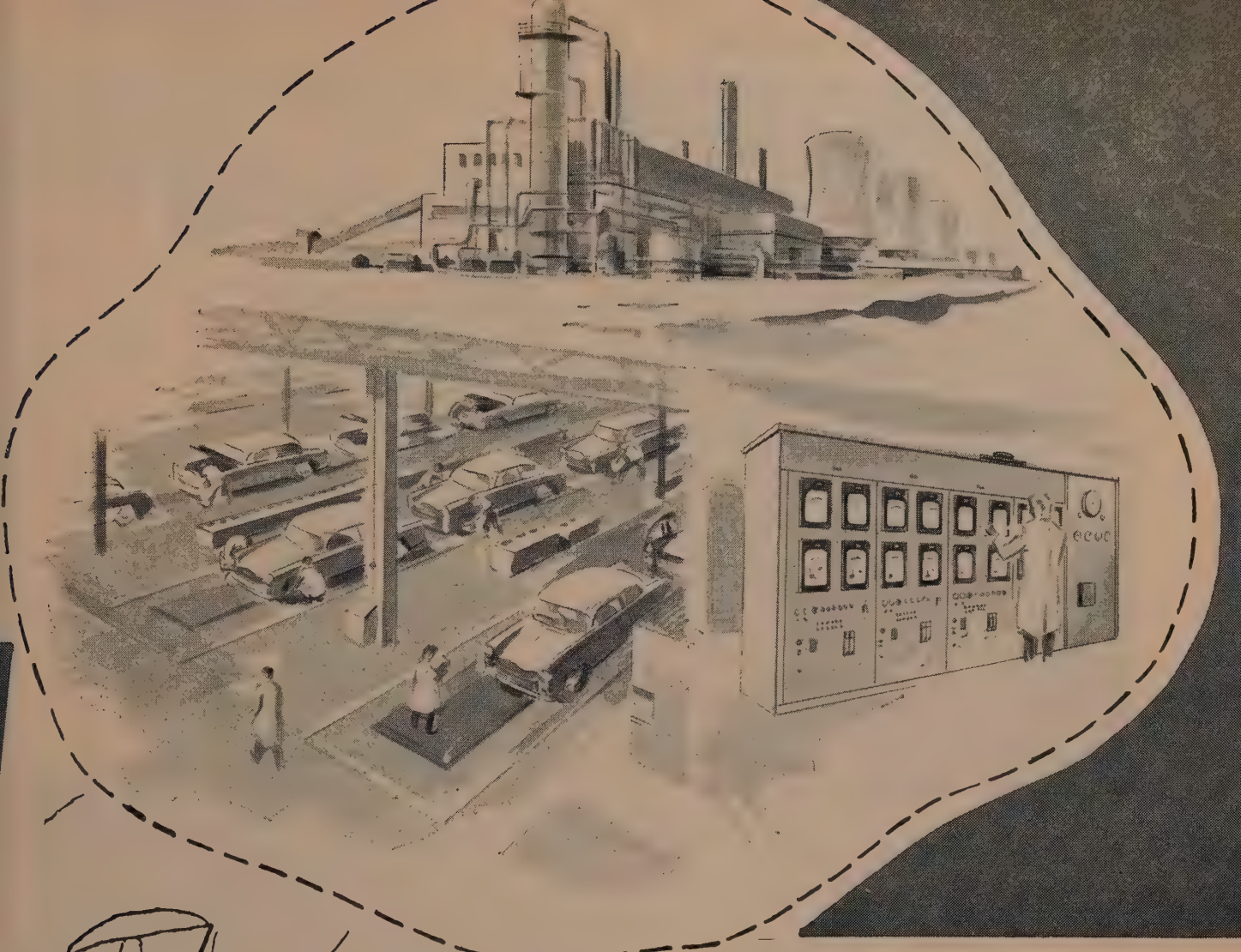
A development of considerable importance in Portuguese telecommunications is the new microwave radio-telephone system between Lisbon and Montijo. This link, supplied to the order of the Anglo-Portuguese Telephone Company Limited connects telephone exchanges on the south side of the Tagus with the main Metropolitan automatic network. Initially it provides 60 telephone circuits to supplement an existing submarine cable connection, but at a later stage the capacity will be increased to 240 circuits. The radio equipment, operating in the 2000 mc/s band has been supplied by Marconi's Wireless Telegraph Co. Ltd., and the associated channelling and signalling equipment by Automatic Telephone & Electric Co. Ltd.



AUTOMATIC TELEPHONE & ELECTRIC COMPANY LTD.
STROWGER HOUSE, ARUNDEL STREET, LONDON, W.C.2.

MARCONI'S WIRELESS TELEGRAPH COMPANY LTD.
MARCONI HOUSE, CHELMSFORD, ESSEX, ENGLAND.

AT 14861



WHAT is BI-stat CONTROL?

BI-stat is the new static control system developed by Brookhirst Igranic. In this system, the most complex information is infallibly sorted and transmitted by small, light, logical units—*entirely devoid of contacts and moving parts.* Nothing to break or burn out! And encapsulated housings ensure complete protection from dirt, moisture or fumes. BI-stat equipment introduces new conceptions of efficiency, reliability and long life. It can be used to great advantage in motor control schemes, replacing conventional relays and other auxiliary switching devices, particularly for installations where processes are complicated, stoppage costly or maintenance difficult. Find out how BI-stat may help you to solve tomorrow's control problems today. Write for leaflet 1809 giving full details.

BROOKHIRST IGRANIC LIMITED

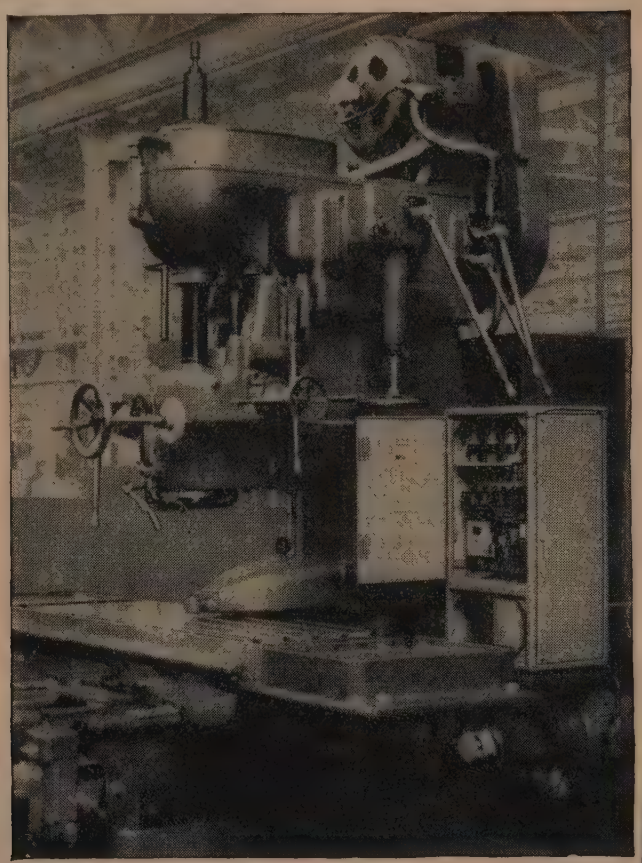
Sales Headquarters: IGRANIC WORKS BEDFORD

HEAD OFFICE & EXPORT SALES DIVISION : BROOK HOUSE PARK LANE LONDON W1

AREA OFFICES: BIRMINGHAM · BRISTOL · CARDIFF · EAST ANGLIA · GLASGOW · LEEDS · LONDON · MANCHESTER MID. SOUTH · NEWCASTLE · NOTTINGHAM · SHEFFIELD

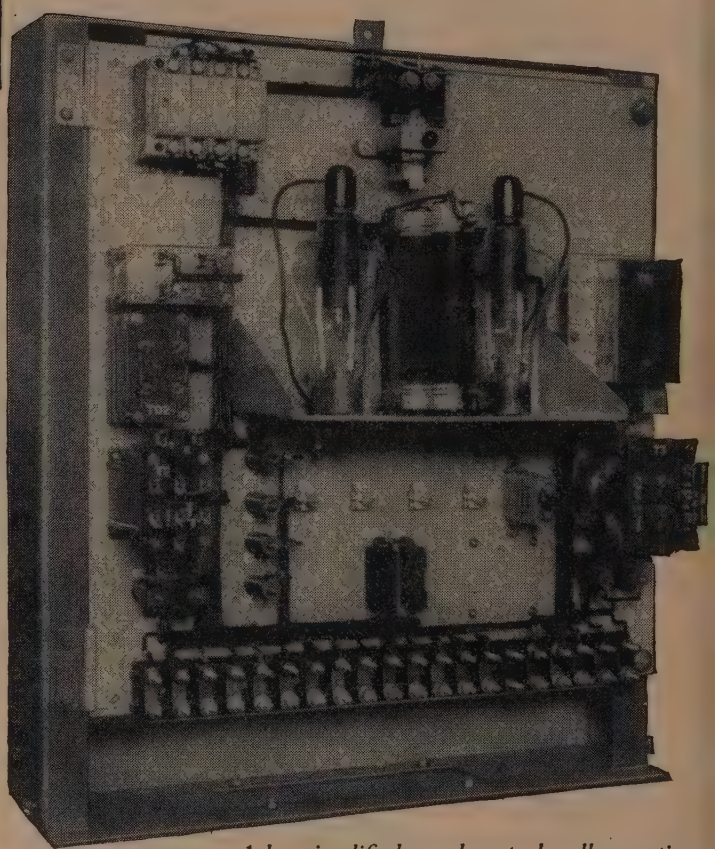
WORKS AT BEDFORD & CHESTER

A Metal Industries Group Company



Metrovick *Quality* Speed Controls

(Left) Herbert 28A vertical milling machine with electronically controlled motor driving the feed. Speed range of 60/1200/3000 rpm.



1 hp. simplified speed control wall mounting.

The reliability of Metrovick electronic motor speed controls has been proved by years of practical service in all branches of industry. The types of speed control available are :—

SIMPLIFIED THYRATRON SPEED CONTROLS
Using a magnetic amplifier and two thyratrons, these equipments offer fingertip stepless speed control of a D.C. motor from zero to top speed. Speed holding accuracy of 2% despite load changes and automatic acceleration to any set speed. This control has a constant torque characteristic, but constant horsepower over a 3/1 range can be obtained.

HIGH ACCURACY THYRATRON SPEED CONTROLS
Using an amplifier with two thermionic valves, two thyratrons and a tachogenerator, these high accuracy controls embody all the features of the simplified equipments with the additional advantage of holding the speed to an accuracy of $\pm \frac{1}{4}\%$ of top speed despite load and supply variations.

MAGNETIC AMPLIFIER CONTROLS
Speed control of motors below 1 hp. Accuracy of speed holding to within 2% of top speed.

For further details write for publications Nos. 98/1-1 7098/3.

METROPOLITAN-VICKERS
ELECTRICAL CO LTD · TRAFFORD PARK · MANCHESTER, 17

An A.E.I. Company

PRACTICAL ELECTRONIC MOTOR SPEED CONTROLS

H/R602

A new service for equipment designers and engineers



VOLS. 1 & 2
SPECIAL VALVES

VOL. 3
TRANSISTORS
MAGNETIC MATERIALS
QUARTZ CRYSTALS
RECTIFIERS · CAPACITORS

VOL. 4 *(to be issued in June)*
BRIMAR VALVES

S.T.C. Components Handbook

quipment designers and engineers will find extensive use for this new service y S.T.C. Data Sheets on the full range of S.T.C. Components are included in nese three loose-leaf volumes and amendments and additions will be made at uarterly intervals. The cost to subscribers of £1 per annum will include the inders, data sheets and all subsequent amendments.

** This applies in U.K. and N.Ireland only. Overseas Subscribers are requested not to send fee with application.*

*Please register me as a subscriber for the S.T.C. Components Handbook. A Cheque/ P.O. for £1 is enclosed.**

Name.....

Company.....

Address.....

.....

I.E.E.B

COMPONENTS GROUP

Standard Telephones and Cables Limited

CONNAUGHT HOUSE · ALDWYCH · LONDON WC2

Marconi in Radar



29 Countries use

Marconi Radar

MARCONI

COMPLETE CIVIL AND MILITARY
RADAR INSTALLATIONS

MARCONI'S WIRELESS TELEGRAPH COMPANY LIMITED, CHELMSFORD, ESSEX, ENGLAND

M 68

THE 5867 TRANSMITTING TRIODE



The 5867 is a transmitting triode with a maximum anode dissipation of 250W. It can be operated at full ratings up to 100 Mc/s, and at reduced ratings up to 150 Mc/s. Natural cooling is normally sufficient for frequencies up to 30 Mc/s, but at higher frequencies it may be necessary to direct one air flow on to the anode seal and another on to the base and thence over the bulb. For full electrical and mechanical details and for tables of typical operating conditions, please write to the Company. Details of any of our large range of valves will be sent on request.

EQUIVALENT TYPES:—
TY3-250
TB3/750
CV1350

This valve is eminently suitable for RF heating, both dielectric and induction.

'ENGLISH ELECTRIC'

ENGLISH ELECTRIC VALVE CO. LTD. *Chelmsford, England*
Telephone: Chelmsford 3491

AP/126



Sticking our necks out?

G.E.C. announce **maximum junction temperature** uprating to **85°C** for their audio transistors.
We know this claim isn't too tall because we have sound reasons for making it.
Recently our manufacturing techniques have been improved to the extent that life tests show that we can now quote an 85°C **continuous working** maximum junction temperature for G.E.C. germanium audio transistors.
As a result of this, the already high maximum collector dissipation ratings have been increased even further. These new ratings, coupled with the typical alpha cut-off frequencies of about 1Mc/s, make the G.E.C. range of audio transistors unique.



G.E.C. AUDIO TRANSISTORS



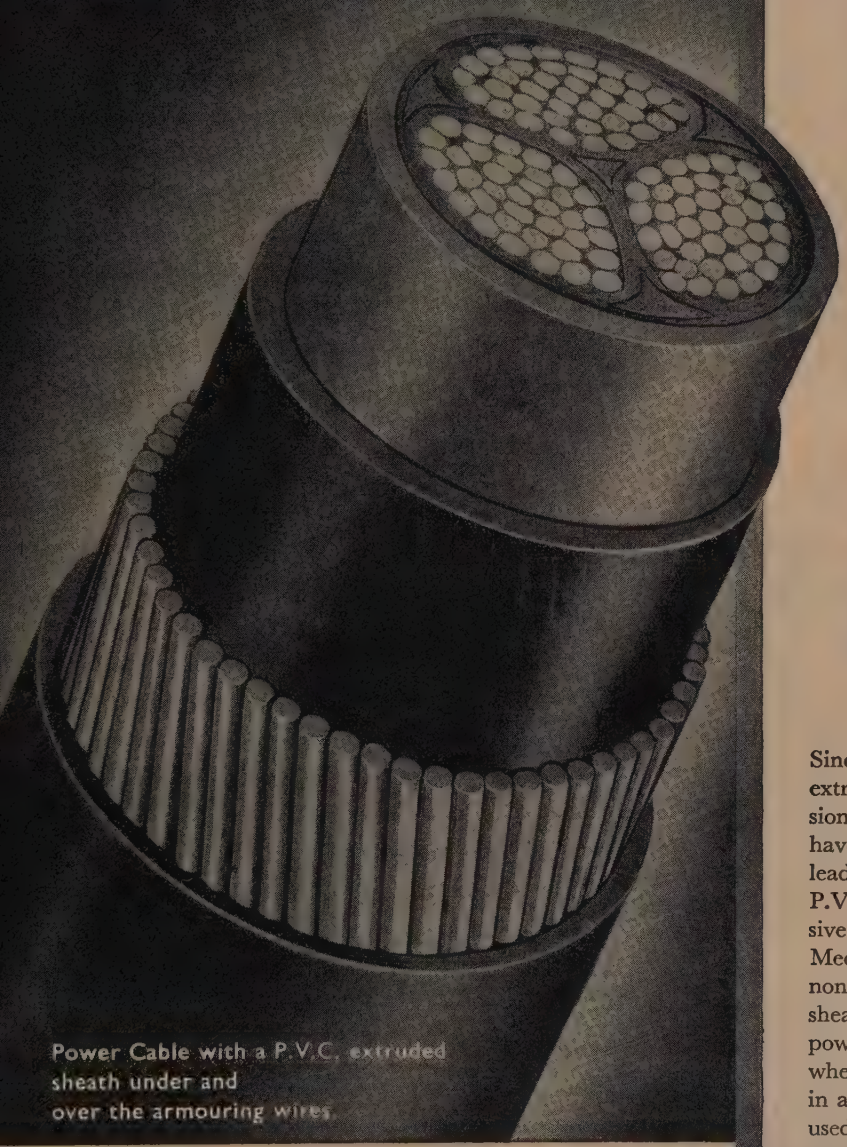
	LOW NOISE	LOW POWER†					MEDIUM POWER†		
	Maximum noise factor=5dB (f=1kc/s, R_s=500 Ω V_{ce}=−2V, I_e=0.5mA)	Maximum collector dissipation at 45°C=200mW at 55°C=150mW					Maximum collector dissipation at 45°C=800mW at 55°C=600mW (on 3″×3″ cooling fin)		
$i_{c(pk)}$(A)	—	1					1		
	GET106	GET114	GET113‡	GET103	GET102‡	GET104	GET115	GET116	GET105
$V_{ce(pk)}$(V)	15	15	15	30*	30*	30	15	30*	40*

‡New high gain types *R_e/R_b>0.03 †Can be supplied in matched pairs

G.E.C. Semiconductor Division, School St. Hazel Grove, Stockport, Cheshire. Tel: Stepping Hill 3811
London enquiries Tem. Bar 8000. Ext. 10

Corrosion protection...

SIEMENS EDISWAN



Power Cable with a P.V.C. extruded sheath under and over the armouring wires.

P.V.C. EXTRUDED SHEATH CABLES

Since we pioneered the use of P.V.C. extruded sheath as a means of corrosion protection several years ago, we have laid and jointed many miles of lead covered cable sheathed with P.V.C. in situations of a highly corrosive nature.

Mechanically strong, flexible and non-inflammable, P.V.C. extruded sheath is suitable for all types of power and communication cables, whether underground, underwater or in air. Bituminous compound can be used under the P.V.C. if required as a safeguard against corrosion by differential aeration.

Suitable for all cable diameters, *armoured or unarmoured*, and for all working voltages.

All enquiries to:

SIEMENS EDISON SWAN LTD. *An A.E.I. Company*
Power Cables Division, 155 Charing Cross Road, London, W.C.2

SIEMENS EDISWAN

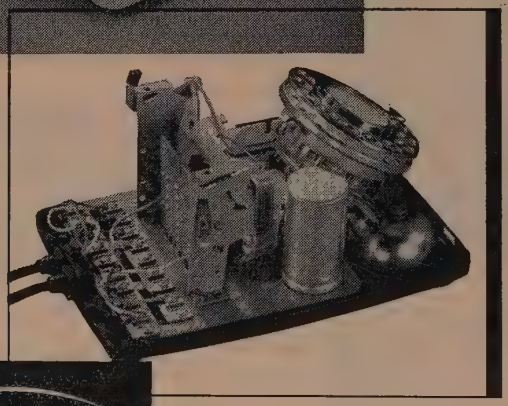
a power in cables

The new British Post Office Telephone



but fully tropicalised





SIEMENS EDISWAN

Anticipating the demand from many overseas countries for the new British Post Office telephone, Siemens Edison Swan have developed a model suitable for use in any climate. This instrument is in full tropical finish, ventilated and effectively sealed against dust and insects. Available in colours.

It can be supplied with or without automatic transmission regulator and an attenuator for the transmitter is also an optional feature. When the automatic regulator is not supplied, provision is made for adding it at any time.

For full particulars write to:

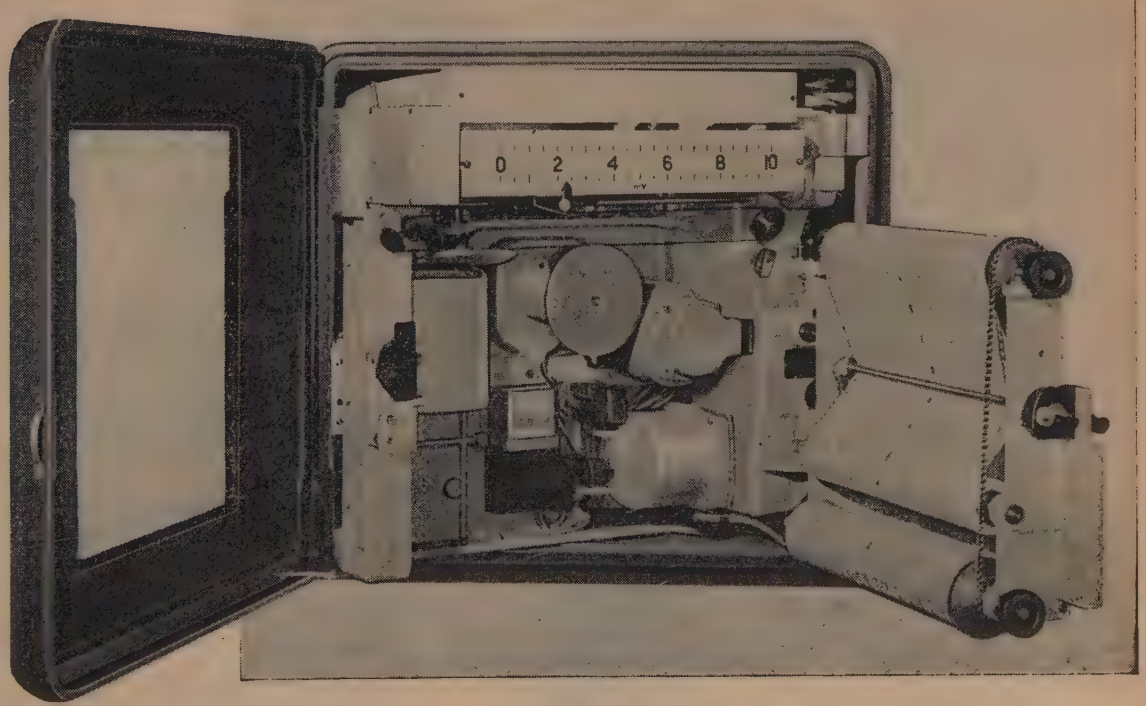
SIEMENS EDISON SWAN LTD.
A.E.I. Company
Telecommunications Division P.D.7/4
Woolwich, London, S.E.18
Telephone: Woolwich 2020 Ext. 724

Announcing the *mark 3* strip-chart electronic *multelec* recorder



Features include

- *Electronic system and components as used in other* KENT *instruments*
- *All-mains operation—no batteries—with continuous standardization*
- *Availability in 2-, 3-, 4-, 6-, 8-, 12- and 16-point form*
- *2-second full-scale travel—deadbeat balancing eliminating "overshoot"*
- *Chart-speed variable between 1 in./h. and 4 in./min.*
- *Multi-point printing intervals of 5, 10 or 20 seconds, depending on application*
- *High-gain amplifier, with extremely sensitive synchronous converter*



Mark 3 Multelec recorder, showing the chart frame swung away to allow ready access to the main-frame components; plug-in amplifier etc. is easily reached by swinging forward the main frame.

KENT masters of instrumentation

Write for Publication 295, or telephone General Industrial Contracts Dept. (Luton 2440, Ext. 8) for a quotation.

GEORGE KENT LIMITED · LUTON · BEDFORDSHIRE · ENGLAND

Factories, Subsidiary Companies, Branch Offices, and Representatives in most countries.

PHILIPS

In Science and Industry alike . . .



among technicians, manufacturers and those engaged in the sale of electrical products — as well as among the public at large, the Philips emblem is accepted throughout the World as a symbol of quality and dependability.

PHILIPS ELECTRICAL LTD

Century House · Shaftesbury Avenue · London · WC2

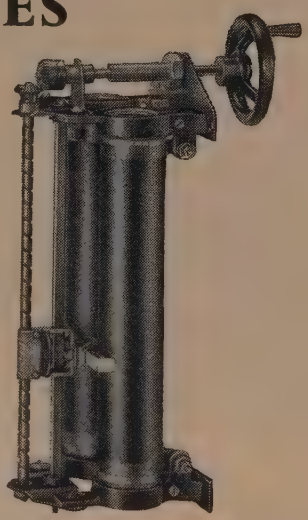
Radio & Television Receivers · Radiograms & Record Players · Gramophone Records · Tungsten, Fluorescent, Blended and Discharge Lamps & Lighting Equipment · 'Philishave' Electric Dry Shavers · 'Photoflux' Flashbulbs · High Frequency Heating Generators · X-ray Equipment for all purposes · Electro-Medical Apparatus · Heat Therapy Apparatus Arc & Resistance Welding Plant and Electrodes · Electronic Measuring Instruments · Magnetic Filters · Battery Chargers and Rectifiers · Sound Amplifying Installations · Cinema Projectors · Tape Recorders · Health Lamps · Hearing Aids · Electrically Heated Blankets

ZENITH

(REGD. TRADE-MARK)

TUBULAR SLIDING RESISTANCES

Back-of-Board Types



Rugged, robust, reliable rheostats, ratings from 200 to 2,000 watts. Clutch fitted to prevent over turning. Laminated phosphor bronze contacts fitted as standard to avoid residual resistance at all out position. Carbon contacts available when desired.

Illustrated catalogue of all types free on request

The ZENITH ELECTRIC CO. Ltd.

ZENITH WORKS, VILLIERS ROAD, WILLESDEN GREEN
LONDON, N.W.2

Telephone: WILlesden 6581–5 *Telegrams:* Voltaohm, Norphone, London

MANUFACTURERS OF ELECTRICAL EQUIPMENT
INCLUDING RADIO AND TELEVISION COMPONENTS

The PRECISION WIRE-WOUND Resistor

—most in demand



Other products include Vitreous coated Silicone and Glass Bond Resistors, Transformers, Chokes & Interleaved Coils.

ERG

Trade Mark

ERG INDUSTRIAL CORPORATION LTD
18 MANCHESTER STREET · LONDON · W.1
Tel: WELbeck 8114/5



35 YEARS OF 'KNOW HOW' BEHIND EVERY TERMINAL

We make millions of terminals in a most comprehensive range of types and sizes. In many cases, collar assemblies are interchangeable, thus making possible a wide variety of combinations. Where our range—extensive as it is—does not provide for your particular requirements, we are always prepared to 'make to measure'.

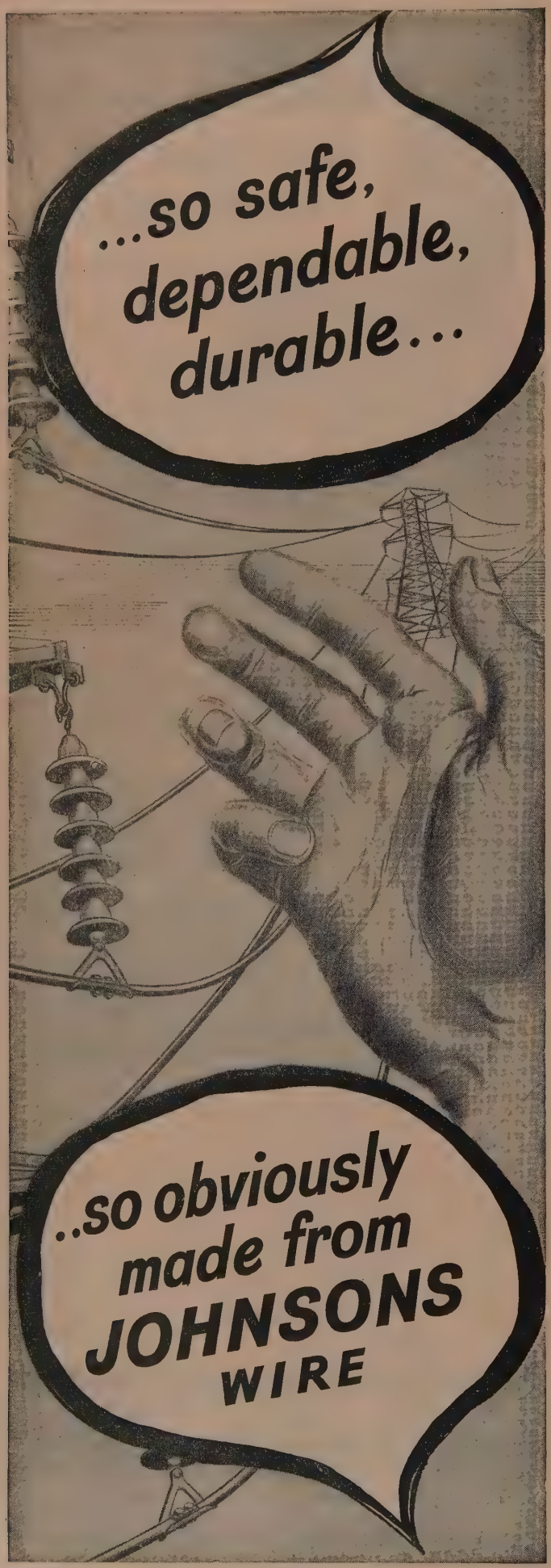
We are specialists in components. This is no empty phrase: it means that we have been engaged in the design and manufacture of precision-made components for some thirty-five years, and at Enfield we devote a very great deal of our time and space to research, forever seeking new and better methods of production, whilst ensuring that every component maintains our high standard of performance.

Most "Belling-Lee" products are covered by patents or registered designs or applications therefor.

BELLING & LEE LTD
GREAT CAMBRIDGE ROAD, ENFIELD, MIDDX., ENGLAND
Telephone: Enfield 3322 · Telegrams: Radiobel, Enfield

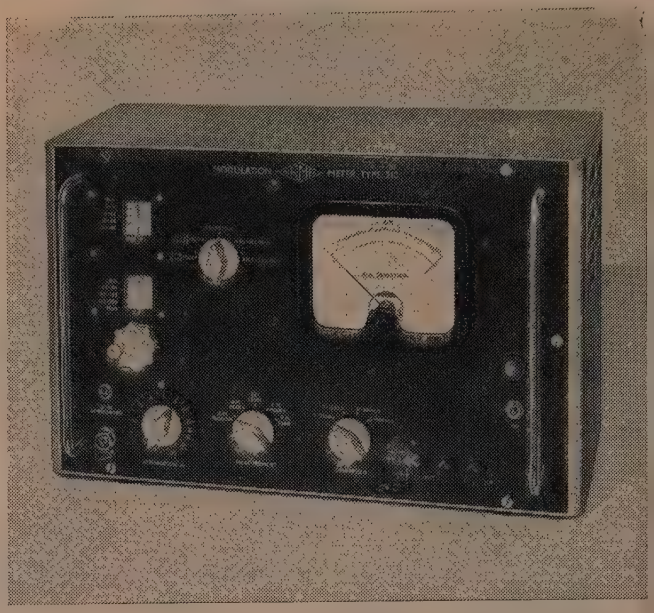
Regd.

TERMINALS · PLUGS & SOCKETS · THERMAL DEVICES
FUSES · INTERFERENCE FILTERS · RECEIVING AERIALS



RICHARD JOHNSON & NEPHEW LIMITED, MANCHESTER 11 Tel. EAST 1431

Measuring Modulation?



AIRMEC **MODULATION METER**

measures

- *percentage modulation of amplitude modulated signals*
- *peak deviation of frequency modulated signals*

THE MODULATION METER Type 210 may be used to measure the percentage modulation of amplitude modulated signals and the peak deviation of frequency modulated signals in the carrier frequency range 2·25 Mc/s to 300 Mc/s.

Outputs at the intermediate frequency of 750 kc/s and at low frequency are available from terminals on the front panel. These outputs enable the modulated envelope of the input signal and the demodulated signals to be observed on an oscilloscope.

The limiting action of the instrument is so effective that it can be used to measure spurious frequency modulation occurring on amplitude modulated signals. Furthermore, changes of mean carrier level when amplitude modulation is applied can be measured to an accuracy of better than ±1%.

Simplicity of Operation

One of the most attractive features of the instrument is its simplicity of operation. The tuning control is adjusted until a meter reading is obtained, and the input attenuator adjusted for full scale deflection. It is then only necessary to switch to the A.M. or one of the F.M. positions to obtain a direct reading of modulation.

Statistics of Airmec Modulation Meter

*	Frequency range . . .	2·25 to 300 Mc/s in 7 bands (up to 600 Mc/s with reduced sensitivity)
*	Input level . . .	A.M. 7–700 mV F.M. 7 mV–10 V (3 kc/s–100 kc/s) 50 mV–10 V (0–3 kc/s)
*	Modulation Frequency range	30 c/s–15 kc/s
*	A.M. range	0–100% ±3%
*	F.M. range	0–100 kc/s ±5%

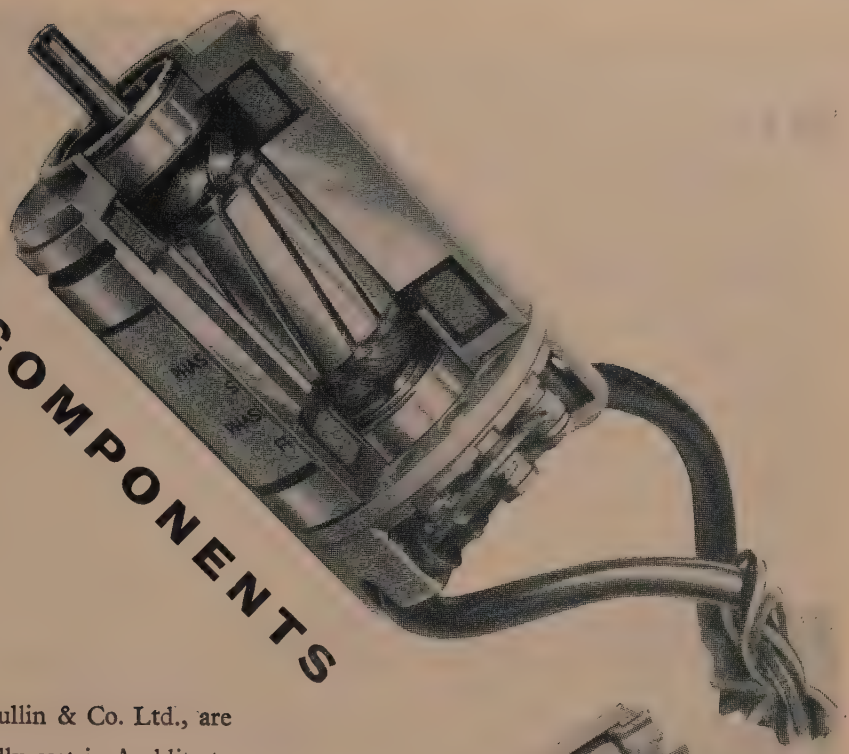
DELIVERY—EX STOCK

Airmec **AM/FM Modulation Meter Type 210**

AIRMEC LIMITED · HIGH WYCOMBE · BUCKS

Telephone: High Wycombe 2060

Araldite IN SERVOCOMPONENTS





The Synchro units shown, made by R. B. Pullin & Co. Ltd., are sectioned to show how the stators are integrally cast in Araldite to provide maximum protection against the effects of extremes of temperature, humidity and vibration. The excellent machining properties of Araldite make possible a straight-through bore technique, which eliminates errors in alignment and also permits the use of the smallest possible air gap between rotor and stator. High insulation and dielectric strength, remarkable adhesion to metals, and negligible shrinkage on curing make Araldite eminently suitable for use in the construction of precision electrical equipment.

Araldite epoxy resins are used—

for casting high grade solid electrical insulation

for impregnating, potting or sealing electrical windings and components

for producing glass fibre laminates

for producing patterns, models, jigs and tools

as fillers for sheet metal work

as protective coatings for metal, wood and ceramic surfaces

for bonding metals, ceramics, etc.



This photograph shows an A.E.W. electric oven, capable of maintaining temperatures within close limits, as used by R. B. Pullin & Co. Ltd. for curing the Araldite-filled stators.

Araldite *epoxy resins*

Araldite is a registered trade name

CIBA (A.R.L.) LIMITED

Duxford, Cambridge Telephone: Sawston 2121

AP 463

Wire for electrical progress



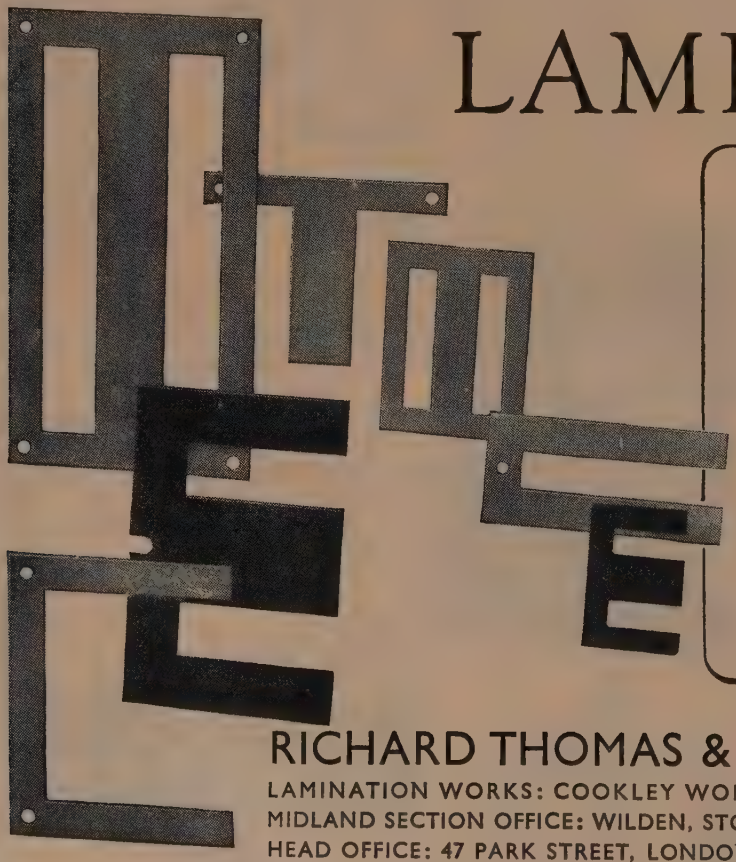
Lewmex or Lewkanex, Lewcosol or Lewcoglass—just a few of the 'LEWCOS' range of insulated wires designed to meet every need or application in the electrical industry, backed by intensive research and a first class technical service.

LEWCOS *the largest manufacturers of insulated wires and strips in Europe*

THE LONDON ELECTRIC WIRE COMPANY AND SMITHS LIMITED

LEYTON · LONDON · E10

LAMINATIONS



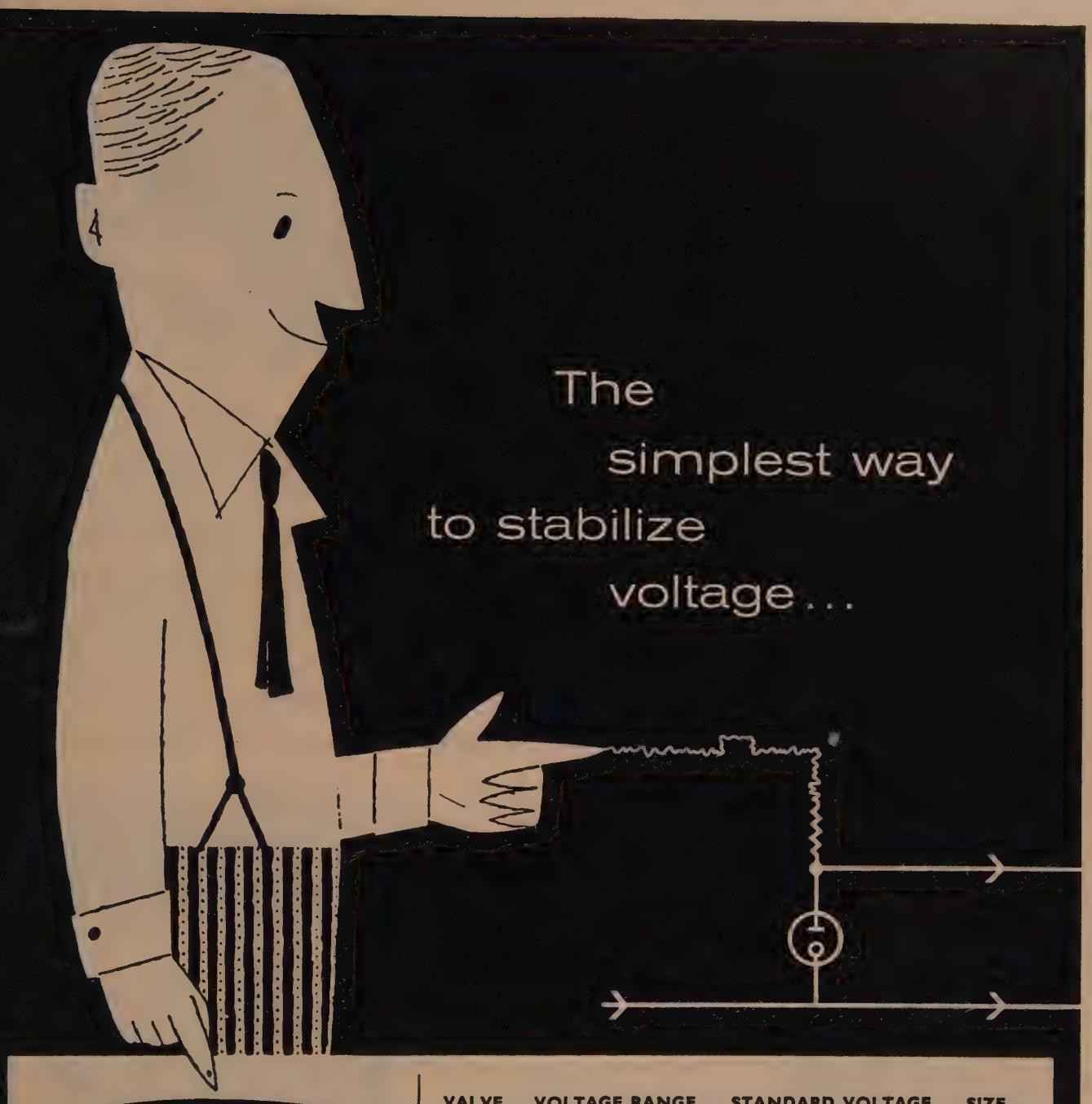
Laminations of all types, in all sizes and in all grades of material

FERROSIL
hot-rolled and cold-reduced electrical sheet and strip, and hot-rolled transformer sheet

ALPHASIL
cold-reduced oriented transformer sheet and strip

RICHARD THOMAS & BALDWINS LTD
LAMINATION WORKS: COOKLEY WORKS, BRIERLEY HILL, STAFFS.
MIDLAND SECTION OFFICE: WILDEN, STOURPORT-ON-SEVERN, WORCS.
HEAD OFFICE: 47 PARK STREET, LONDON, W.1

Our Cookley Works is one of the largest in Europe specializing in the manufacture of laminations for the electrical industry.



G.E.C.

CORONA STABILIZERS 350V to 7kV

VALVE	VOLTAGE RANGE	STANDARD VOLTAGE	SIZE
SC1	350-2000V	350, 400, 600, 800, 1000 1200, 1400, 1600, 1800, 2000	B7G
SC2	2000-4500V	2500, 3000, 3500, 4000	B9A
SC4	4500-7000V	5000, 6000, 7000	B9A
SC3	350-400V	350, 400 (Low current)	1½" + Flying leads

Minimum currents range from 5μA to 50μA, maximum currents range from 50μA to 2mA, according to type.

Please write for further information to—

The M-O Valve Co. Ltd., Brook Green, Hammersmith, London, W.6

Makers of G.E.C. high grade radio valves and cathode ray tubes
A subsidiary of The General Electric Co. Ltd.

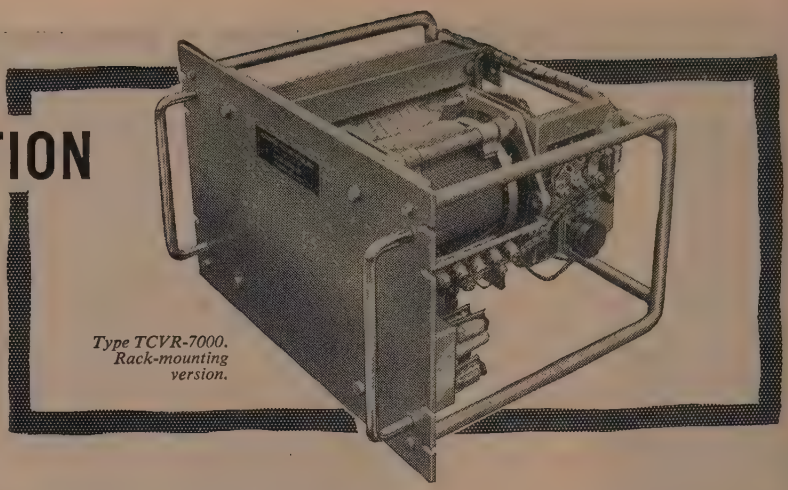
40 VOLTS/SEC AUTOMATIC CORRECTION

—with the type TCVR voltage regulator

The TCVR is a servomechanical automatic voltage regulator having the very high speed of correction of FORTY VOLTS PER SECOND. It provides an *undistorted* output, maintained constant within very close limits (normally ±0.5%) from no-load to full-load, for wide variations in frequency and power factor.

A wide range of models from 1.6 to 12 kVA single-phase, and 4.8 to 36 kVA three-phase, is available, to standard or tropical specification, in cabinets or for rack-mounting. Models are also available in which the output voltage is continuously adjustable over a wide range by means of a panel control. Regulators can be supplied to Services' specification, and special models can be designed to order.

For high-speed, accurate stabilisation *without distortion*—**specify TCVR.**



Type TCVR-7000. Rack-mounting version.

Other products of Claude Lyons Ltd. Stabiliser Division

BMVR: Motor-driven laboratory and industrial regulators ranging from 1·6 to 29 kVA single-phase, and 4·8 to 87 kVA three-phase. Constancy of output normally ± 0·5%, from no-load to full-load. No distortion. Speed of correction 1 Volt/Sec. A great variety of models, standard, tropical and militarised, for all applications.

BAVR: Electronic stabilisers of very high accuracy, and very rapid response, with no moving parts. Input range: − 10% to + 5%, output constancy ± 0.15%. Three sizes: 200, 500 & 1000 VA. Exceptionally useful for control of chemical processes, heating, lighting, etc.

ASR: Automatic step regulators, small, inexpensive, and with sinusoidal output waveform. Two sizes: 1·15 kVA and 2·3 kVA. Input range − 10% to + 5%; output constancy, ±2½%.

ATC: Automatic Tap-Changing Transformers — a development of ASR. Two sizes: 575 VA and 1150 VA. Input range − 20% to + 10%: output constancy, ± 5%. Provide adequate stabilisation for many types of apparatus, at low cost. Also useful as pre-regulators, e.g. in conjunction with BAVR.

We shall be pleased to send you full details of our entire range.

Claude Lyons Ltd.

Stabiliser Division **VALLEY WORKS · HODDESDON · HERTS · TELEPHONE HODdesdon 3007-8-9**

CL/47/E2B

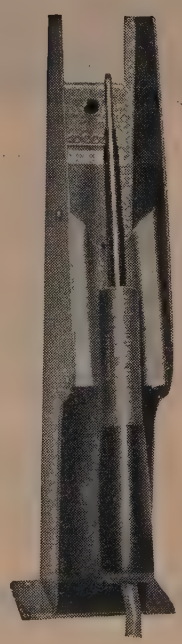
ADCOLA
(Regd. Trade Mark)

Soldering Instruments

ILLUSTRATED
PROTECTIVE SHIELD (CAT. No. 68)
⅛ in. BIT MODEL (CAT. No. 70)

Primarily developed for the TRANSISTOR & ELECTRONIC ERA.

Possessing the sharp heat essential for the quick jointing of Transistors, Resistors, etc., thereby avoiding damage to the equipment from heat transference



Cover all requirements for thorough solder jointing in all the fields of

TELECOMMUNICATIONS

Fully Insulated Elements

Suited to daily use for bench line production

MANUFACTURED IN ALL VOLT RANGES

British and Foreign Pats. Reg. designs, etc.

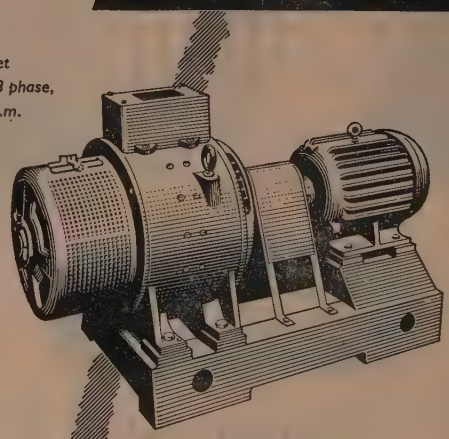
For further information apply Head Office:

ADCOLA PRODUCTS LTD.
GAUDEN ROAD
CLAPHAM
HIGH STREET
LONDON, S.W.4

Tel.: MACaulay 4272 & 3101

ELECTRICAL EQUIPMENT

High Frequency Motor Alternator Set 5 kVA, 208 volts, 3 phase, 400 c.p.s. 3000 r.p.m.



A leaflet giving full particulars will be sent on request.

NEWTON DERBY

NEWTON BROS. (DERBY) LTD

ALFRETON ROAD, DERBY

PHONE: DERBY 47676 (4 LINES) *GRAMS:* DYNAMO, DERBY
London Office: IMPERIAL BUILDINGS, 56 KINGSWAY W.C 2

High grade transmission systems

Radio Links 5-12-24-48-60-120 Circuits

Radio Telephone Terminals with optional channel shifting, privacy and control equipment

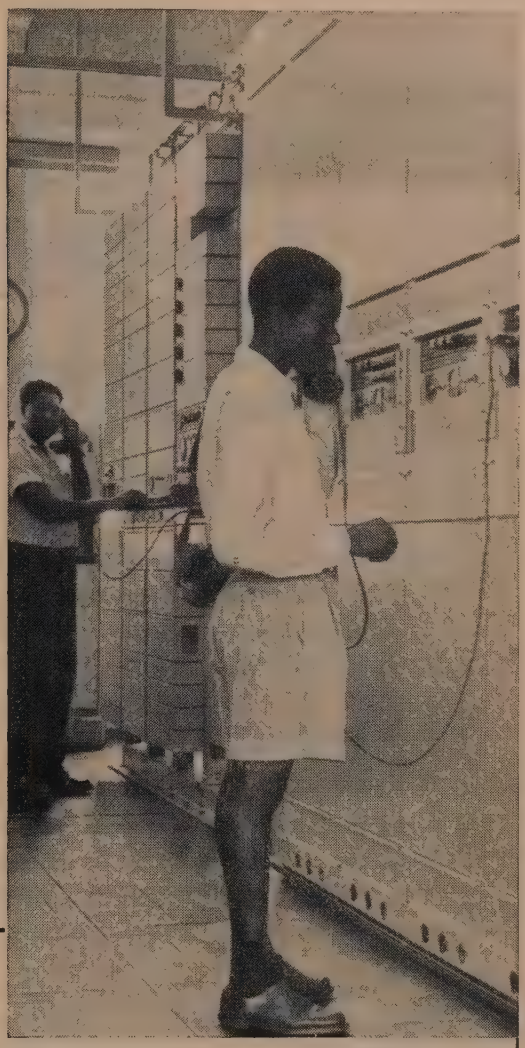
Carrier on cable telephone systems

Carrier programme channelling equipment

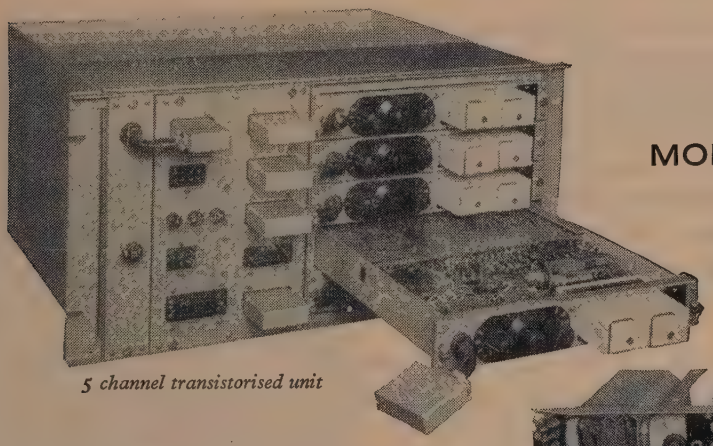
Submerged repeater systems

High speed FMVFT equipment

Portable carrier systems for military, police and civilian use



Typical overseas terminal



5 channel transistorised unit

MODERN CIRCUIT TECHNIQUES

Transistors Resin Cast Components
Printed Wiring Crystal Frequency Control
Out of Band Signalling (E & M or Ringdown)

Transistorised channel unit



We shall be pleased to advise and provide full information upon request.

extending → **the frontiers of telecommunications**

SIEMENS EDISWAN

SIEMENS EDISON SWAN LTD *An A.E.I. Company*
Telecommunications Division P.D.8, Woolwich, London SE18, England
Cables: Sieswan London

D8/5

TAS 2313

Living Together

Pluggable Components

THE MOST RECENT, the most outstanding, and the most revolutionary example which has emerged from the Erie principle of living together, and one which will be welcomed by all associated with the problems of adapting traditional components for printed circuits, is a component specially developed and specially tailored for the job.

The Erie pluggable component is fitted with special strip terminations, shouldered and tapered for easy insertion and positive location, thus avoiding looping, crimping, bending, cropping, and elaborate and expensive insertion machinery, as is necessary with the traditional wire ended component.

The design of the termination eliminates the possibility of the component falling out or loosening prior to the soldering operation, the finish of the termination ensures consistently good connection with the minimum amount of solder, and the shoulder on the termination raises the component to a standard and safe distance from the board. Furthermore, the design of the termination ensures that inductance and stray capacitance are low and constant, and, above all, facilitates replacement in servicing.

Engineers and designers interested in saving costs on printed circuit applications are invited to write for details and samples.



Style AP Ceramicon★



Style BP Ceramicon★



Type 8AP Resistor



Style 811P Ceramicon★

Designs Registered
Patents Pending

1, HEDDON STREET, LONDON, W.1

FACTORIES
Great Yarmouth and Tunbridge Wells, England: Trenton, Ont., Canada: Erie, Pa., Holly Springs, Miss., and Hawthorne, Cal., U.S.A.

★ERIE
RESISTOR
LIMITED

★*Registered Trade Marks*

INDEX OF ADVERTISERS

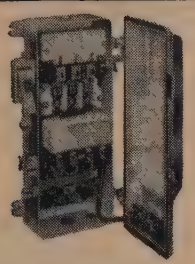
Firm	page
Adcola Products Ltd.	xl
Airmec Ltd.	xxxvi
Sir W. G. Armstrong Whitworth Aircraft Ltd.	v
Automatic Telephone & Electric Co. Ltd.	vii, xxiv + *IBC*
Belling and Lee Ltd.	xxxv
British Thomson-Houston Co. Ltd.	
Brookhirst Igranic Ltd.	xxv
Cable Makers' Association	
Cathodeon Crystals Ltd.	
Ciba (A.R.L.) Ltd.	xxxvii
Dewhurst and Partner Ltd.	
Donovan Electrical Co. Ltd.	xliii
Dubilier Condenser Co. Ltd.	xi
English Electric Co. Ltd.	xxiii
English Electric Valve Co. Ltd.	xxix
Erie Resistor Co. Ltd.	xlii
E.R.G. Industrial Corp. Ltd.	xxxiv
Ferranti Ltd.	ii + iii
P. X. Fox Ltd.	
General Electric Company Ltd. (M.O. Valves)	xxxix
General Electric Company Ltd. (Semi conductors)	xxx
General Electric Company Ltd. (Telecommunications)	vi + vii
R. Johnson and Nephew Ltd.	xxxvi
George Kent Ltd.	xxxiii
Lodge Plugs Ltd.	
London Electric Wire Co. and Smiths Ltd.	xxxviii
Claude Lyons Ltd.	xl
Marconi Instruments Ltd.	xxii
Marconi Wireless Telegraph Ltd./A.T.E.	xxiv
Marconi Wireless Telegraph Ltd.	xiii, xxviii
Metropolitan Plastics Ltd.	
Metropolitan-Vickers Electrical Co. Ltd.	xxvi
Modern Acoustics Ltd.	xviii
Mullard Ltd. (Magnetic)	
Mullard Ltd. (Valves)	xvii
Newmarket Transitor Co. Ltd.	xix
Newton Bros. (Derby) Ltd.	xl
Oliver Pell Control Ltd.	
Philips Electrical Ltd.	xxxiv
Plessey (Swindon) Ltd.	
Racal Engineering Ltd.	vi
Rank Cintel Ltd.	iv
Redifon Ltd.	i
Richard Thomas & Baldwins Ltd.	xxxviii
Salford Electrical Instruments Ltd.	
Savage Transformers Ltd.	
Semiconductors Ltd.	xii
Servomex Controls Ltd.	
Siemens Edison Swan (Cables)	xxxi
Siemens Edison Swan (Public Telephones)	xxxii, xli
Siemens Edison Swan Ltd. (with Redifon)	i
Siemens Edison Swan Ltd. (Telecommunications)	xvi
Siemens Edison Swan	
Smiths Industrial Instruments Ltd.	
Standard Telephones and Cables Ltd.	x, xiv, xv, xxvii
Texas Instruments Ltd.	xliv
Telephone Manufacturing Co. Ltd.	xx, xxi
Zenith Electric Co. Ltd.	xxxiv

DONOVAN

SPECIALISTS IN INDUSTRIAL CONTACTOR GEAR, AND ALLIED EQUIPMENT

A FEW EXAMPLES FROM A WIDE RANGE

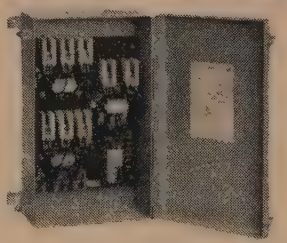
A.C. AND D.C. STARTERS



A.C. DIRECT-ON-LINE CONTACTOR with load breaking interlocked isolating switch. Up to 30 h.p.



A.C. DIRECT-ON-LINE CONTACTOR STARTERS Sizes up to 400 h.p.



Automatic STAR-DELTA STARTERS Up to 150 h.p.

SPECIAL PURPOSE CONTROL PANELS



FOR MOST INDUSTRIES AND APPLICATIONS

Automatic Gear can give almost any required operation and while some applications would not be economic, a surprisingly large number to-day are not only possible but extremely profitable. Our Automation Engineers will be glad to advise you.



LIMIT SWITCHES for A.C. and D.C. control circuits.

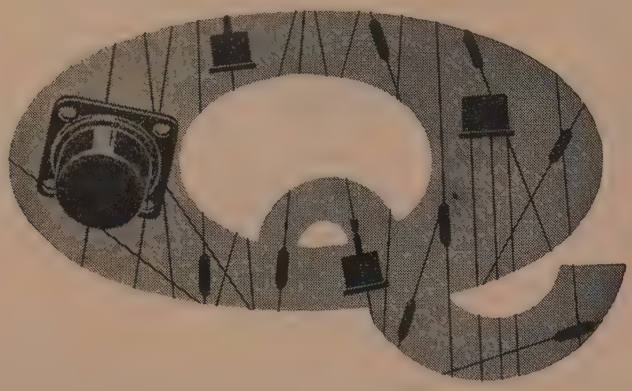


A wide range of PUSH-BUTTONS for ordinary or heavy duty.

CONTROL ACCESSORIES

The DONOVAN ELECTRICAL CO. LTD. 70-81 GRANVILLE STREET, BIRMINGHAM 1.

LONDON DEPOT: 149-151, York Way, N.7. GLASGOW DEPOT: 22, Pitt Street, C.2.

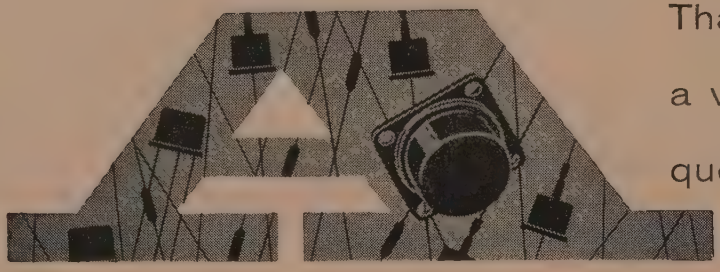


How much do

TEXAS

TRANSISTORS

cost?



That's

a very good

question

THEY COST LESS...

Less than they did. Texas prices are coming down every day, but we will never, repeat never, let this lower our standards of quality.

That is why Texas transistors have such an outstanding reputation for reliability. As a result, demand is constantly increasing and we are constantly producing more.

The more we make, the less each one costs to make. Hence our new price reduction, which came into effect on January 1.

Incidentally, it is the third reduction we have been able to make in twelve months—carrying out our policy of producing better devices at lower prices.

SMALL SIGNAL TRANSISTORS · MEDIUM POWER TRANSISTORS
HIGH POWER TRANSISTORS · SILICON RECTIFIERS

If you have not yet had our new price list we shall be pleased to send you a copy.

Visit us on Stand No. 6 at the I.E.E. Semiconductor Exhibition, Earls Court, 21st to 28th May.

TEXAS INSTRUMENTS LIMITED

Telephone: BEDFORD 68051. **DALLAS ROAD, BEDFORD** *Cables:* TEXINLIM, BEDFORD

T 26

The Institution is not, as a body, responsible for the opinions expressed by individual authors or speakers.
An example of the preferred form of bibliographical references will be found beneath the list of contents.

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 106. PART B. NO. 27. MAY 1959

621.396.713 : 534.84

The Institution of Electrical Engineers
Paper No. 2676 R
June 1958
©

THE ACOUSTIC DESIGN OF TALKS STUDIOS AND LISTENING ROOMS

By C. L. S. GILFORD, M.Sc., F.Inst.P., Associate Member.

(*The paper was first received 9th January, and in revised form 21st March, 1958. It was published in June, 1958, and was read before the* RADIO AND TELECOMMUNICATION SECTION *15th December, 1958.*)

SUMMARY

The paper gives the current views of the author and his colleagues in the Engineering Department of the British Broadcasting Corporation on the design and construction of talks studios and listening rooms or control cubicles, which are considered together on account of their similarity with respect to acoustic behaviour. It is shown that a distinctive characteristic is that, because their dimensions are comparable with the wavelength of low-frequency sound, the sound field is characterized by strong simple standing-wave patterns which cannot be eliminated without eliminating the reverberation itself. It is shown how so that the audible effects are confined to those associated with simple axial modes and that, by careful adjustment of dimensions, provision of diffusion and the proper distribution of absorbing material, the worst faults can be avoided. The effects of the monaural listening chain are considered as well as the consequent necessity for reduced background noise and reverberation in studios as compared with a normal living-room.

Finally, design data for both talks studios and control or listening cubicles are given.

(1) INTRODUCTION

There are over 120 B.B.C. studios used for talks, news, discussions, or continuity announcements. There are also about 50 acoustically treated rooms used for the control and monitoring of programmes, quality checking and similar purposes. All these have one common feature: they are comparatively small, with volumes between 30 and 120 m^3.

In their acoustic behaviour there are two main differences between large and small studios. The most fundamental is that, whereas the wavelength of sounds in the low-frequency end of the audible spectrum is of the order of, or even greater than, the dimensions of a small room, a very large studio has dimensions large compared with all but the very longest relevant wavelengths. Consequently it may be shown that such a small room will have clearly defined resonances characterized by strong standing-wave patterns which are not altered in their general nature by small changes in the room shape, whereas such resonances are usually absent from larger buildings. In place of clearly defined frequency effects, however, large buildings have echoes which are characterized by their time delays and which can be influenced by small changes in the shapes of the walls and ceiling.

The second difference, mainly practical but only slightly less fundamental, is that the reverberation times associated with large studios are generally longer than those of small studios. This is partly because with surfaces of a given average absorption coefficient the reverberation time is proportional to the cube root of the volume, and partly because the optimum reverberation time for a large studio is generally greater than that for a small one. In a small studio the time may be so short that reverberation is no longer appreciable as a time-extension of the original sound, as it is in an orchestral studio or a concert hall, but only as an alteration of the frequency content, making speech, for instance, sound more 'bassy' or more 'sibilant'.

To summarize, large studios are characterized by reverberation and echoes with recognizable time-scales, and small ones mainly by phenomena recognizable as frequency effects. These differences result purely from the disparity in size, and small-studio problems are shared by all small rooms for which good acoustics are required. It seems appropriate, therefore, to consider small studios and listening rooms together in a single paper, first dealing with their common acoustic properties and later describing the detailed treatment required to make them suitable for their different uses. The term 'listening' room' will be used throughout to mean any room such as a quality-checking room, control cubicle or control room which requires good listening conditions but is not used as a studio.

(2) GENERAL ACOUSTIC PROPERTIES OF SMALL ROOMS

(2.1) Formation of Simple Modes in Small Rooms

It has been stated above that the most important characteristic of a small room is the fact that its dimensions lie near or within the wavelength range of low audible frequencies and that this gives rise to recognizable resonance effects known as colorations. If we consider first the simplest possible type of 'mode' or standing-wave system, which is formed when plane waves are reflected between a pair of opposite walls parallel to each other,

Mr. Gilford is with the British Broadcasting Corporation.

1959 : The Institution of Electrical Engineers

the lowest frequency at which a standing-wave pattern will be formed is given by the quotient of the velocity of sound and twice the distance between the walls. At this frequency there will be pressure antinodes at the boundary surfaces and a nodal plane half-way between them. Multiples of this fundamental frequency will give modes with antinodal planes dividing the distance between the boundaries equally into two, three or more equal parts. The frequencies of these modes are given by

$$f_n = \frac{nc}{2l} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where f_n is the nth harmonic of the fundamental, l the relevant room dimension and c the velocity of sound.

There will be room modes of this kind in all parts of the audible spectrum, but whether or not they will be appreciable as colorations depends on the following factors:

(*a*) The bandwidth of the mode.
(*b*) The degree of excitation of the mode.
(*c*) Its separation from neighbouring strongly excited modes.
(*d*) The positions of the sound source and microphone, with respect to standing-wave systems.
(*e*) The frequency content of the source.

Since the colorations largely determine the sound quality from a studio or the goodness of a listening room, the next two Sections will be devoted to an examination of these five factors.

(2.2) Bandwidth of a Room Mode

The ratio of the pressure at a modal frequency f_n to that at any other frequency f may be shown to be

$$\frac{p_{max}}{p_f} = \frac{1 - 2r^2\cos(4\pi fl/c) + r^4}{(1-r^2)\{1 + [1 + \cos(4\pi fl/c)]\}r + r^2} \quad . \quad (2)$$

where r is the mean reflection coefficient of the walls.

The derivation of this equation, which follows without difficulty from the image-plane concept described in Section 2.3, will be omitted for reasons of space.

Defining the bandwidth conventionally as the frequency difference between points on either side of f_n at which the pressure has fallen to $1/\sqrt{2}$ of its peak value, we find:

$$\text{Bandwidth} = \frac{c}{2\pi l}\arccos\frac{(1+r)^4 - (\sqrt{2})(1-r)(1-r^3)}{(\sqrt{2})(1-r)^2 + 2r^2}$$

$$= f_1\phi(r) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where $\phi(r)$ is a function of r alone.

It is clear that the bandwidth depends only on the reflection coefficient of the walls and the frequency of the fundamental mode. In a room having a reverberation time independent of frequency, therefore, all harmonics of a given mode will have substantially the same numerical bandwidth. Such a room will thus have series of modes of comparable bandwidth over the whole audible range, the narrowest bandwidths being associated with the longest room dimension. A typical value for a small studio is about 5 c/s.

(2.3) The Relative Importance of Axial, Tangential and Oblique Modes

So far we have considered only simple 'axial' modes formed by reflection between two parallel wall surfaces. Two other classes are possible: those formed by reflection between two pairs of surfaces, known as 'tangential' modes, and those involving all three pairs of surfaces, known as 'oblique' modes.

The relative subjective importance of the three classes has been dealt with in an important paper by Mayo,[1] who calculated rates of build-up of reverberant sound pressure from individual images and groups of images arranged in lines and planes. Fig. 1 is a two-dimensional representation of the images formed by a point source in one corner of a room, the spacing in the

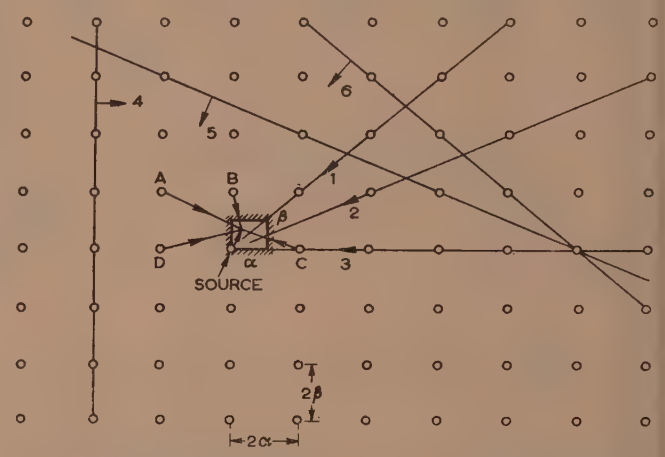


Fig. 1.—Arrangement of images in one plane surrounding a rectangular studio, showing random reflections from images A, B, C, D, line arrays 1, 2, 3, and plane arrays 4, 5, and 6.

two directions being twice the corresponding dimensions of the room.

Mayo considers the way in which the sound pressure at a point in the room builds up when the sound source in the corner is suddenly started. Simultaneously with the start of the sound, all the images appear and start to radiate as point sources, the sound from them reaching the room in order of their distances. Each image gives rise to radiating spherical wavefronts, the pressure diminishing inversely as the distance. Furthermore, there is a loss of pressure by absorption at each reflection, and after the nth reflection the strength of each image has been reduced in the ratio of $r^n : 1$, using the notation of the previous Section. Consideration of the manner in which the images are formed shows that the number of reflections is roughly proportional to the distance from the source, so that the pressure amplitude in the room due to a particular image from which the sound is arriving at time t may be seen to be roughly proportional to r^{kt}/t, where the constant k is a function of the room dimensions and the velocity of sound.

Considering first the sound from the nearest few images, we see that there is no special relationship between their distances, and the wavefronts arrive at random time-spacing and in random phase. The amplitude rises quickly since the images are close, but there is no systematic reinforcement of particular frequencies. Owing to the fact that the pressure decays rapidly according to the same law as that given in the last paragraph, the effect is very temporary and soon gives way to a second regime, involving groups of images arranged in rows.

If we consider the effect of a row of images lying on a line which passes through the studio, we see that the further images of the series lie at distances which increase by equal steps, whereas this is not the case with the earliest images. If these equal steps are multiples of the wavelength, the contributions of the separate images will be in phase and reinforce each other strongly, points in the room lying nearest to the image line attaining the highest pressures.

The characteristic image-row frequencies are given by the formula

$$\tfrac{1}{2}c/\sqrt{[(n_1\alpha)^2 + (n_2\beta)^2 + (n_3\gamma)^2]} \quad . \quad . \quad . \quad (4)$$

here α, β and γ are the dimensions of the room and n_1, n_2 and can be any whole numbers or zero.

As with the random images, however, the decay rate is a oduct of r^{kt} and the inverse of the time. After an initial high nplitude, therefore, the pressure due to these image-row systems llapses quickly and they contribute little to the long-term verberation.

Finally, we have the development of the true room modes hich are formed by the propagation of approximate plane waves om sets of images arranged in planes. To understand the ay in which these modes build up, we remember that each nage constitutes a point source of sound from which spherical aves are propagated. If we consider the wavefront coming om a whole array of images lying in one plane, we see that at a ort distance from the plane the wavefront consists of a set of rtial spheres intersecting at their edges. As we go further vay from the image plane, these spherical surfaces approach arer and nearer to planes until at a great distance they coalesce form a single plane wavefront. The distance, and hence the me, before this process can be regarded as effective is shown by fayo to be much greater than that for the establishment of nage-row frequencies.

However, since the wavefronts are plane and do not diverge, ere is no inverse-distance attenuation, and the rate of decay pends only on the reflection coefficients of the walls. Once tablished, therefore, these true room-modes take a long time die away, and constitute the main reverberant energy in the om.

The frequencies corresponding to these modes are determined the distance between the adjacent planes of images, and are presented by the formula

$$\frac{c}{2}\sqrt{[(n_1/\alpha)^2 + (n_2/\beta)^2 + (n_3/\gamma)^2]} \quad . \quad . \quad . \quad (5)$$

y reference to Fig. 1 it will be seen that the planes with the reatest numbers of images will be those running parallel to the alls of the room. The modes to which they give rise are sually described as 'axial', and since only one dimension of the om is involved, two of the n's in expression (5) are zero and it duces to the simpler form of eqn. (1). Similarly, the tangential odes have one zero n and oblique modes no zero n's.

The individual contributions of the higher harmonics of the mple modes are weak, and those of the high-frequency funda-entals (which are necessarily tangential or oblique) are quite gligible because the images are widely spaced within the plane.

We can now see the way in which the reverberant sound ressure will rise and decay during and after the utterance of a ngle syllable of speech or note of music. At first there will be e reinforcement of the direct sound by randomly spaced reflec-ons reaching a high initial amplitude but decaying sharply after e end of the syllable. The effect will be greatly dependent the positions of the source and microphone. Next, the nage-row frequencies will appear, with high steady intensity but rapid decay. The true modes will meanwhile be building up owly, acquiring only moderate intensity, but decaying slowly ough to provide the bulk of the audible reverberation.

The relative importance of the two series, having regard to eir build-up and decay times and their maximum intensities, pends on the mean reflection coefficients of the walls. For the se of a typical small studio with a mean absorption coefficient about 0·3, calculation shows that no frequency is likely to come prominent unless it is common to both image-row and nage-plane systems, giving a high early intensity and a long cay. This condition is satisfied only by the axial modes, hich are therefore the only ones likely to become individually gnificant. An exception to this rule is that a few tangential or oblique modes of low frequency may possibly be audible, owing to their high initial intensities or wide spacings.

The analysis of a room is therefore considerably simplified since the frequencies of the axial modes form three simple arithmetic series. Calculation of the tangential and oblique modal frequencies is much more laborious, for their total number increases roughly as the cube of the frequency. This is illustrated by Table 1 calculated for an experimental talks studio of typical size.

Table 1

NUMBERS OF AXIAL, TANGENTIAL AND OBLIQUE MODES OF A TYPICAL TALKS STUDIO

Frequency limits	Axial	Tangential	Oblique	Total
c/s				
0–50	2	0	0	2
50–100	3	6	1	10
100–150	4	11	7	22
150–200	3	19	16	38
200–250	3	19	29	51
250–300	4	30	51	85
300–350	2	34	69	105
350–400	3	46	98	147

(2.4) General Conditions for Audibility of Room Modes

Having established that the axial modes alone are likely to have sufficient amplitude and duration, we can now examine the other conditions for their audibility. A mode is heard as a 'coloration' if there is a noticeable tendency for reinforcement at the modal frequency or for sound of neighbouring frequencies to rise or fall in pitch towards that of the mode during the decay time. These processes will be expected to occur if there are, in the frequency region considered, strongly excited modes separated by a frequency interval great in comparison with their bandwidths. Now, it has been shown above that the axial modes form three series with uniform spacing throughout the audible frequency range and that the bandwidths are likely to be about the same for all members of a series. Therefore, we might expect the modes to be equally audible in all parts of the voice-frequency range; in fact, the region of audibility is restricted for other reasons. The most important of these is the presence of tangential and oblique modes, which, though virtually absent at frequencies for which the wavelength is comparable with the room dimensions, are far more numerous than the axial modes in the majority of the voice-frequency range. These non-axial modes, though not individually significant, dissipate an appreciable fraction of the sound energy, and reduce the intensity of the axial modes to such an extent that they are no longer prominent.

This was clearly brought out by some experiments on artificial reverberation[2] in which sound was prolonged by being repeatedly made to traverse a long tube. As the reverberation obtained from such a system exhibited strong resonances, the effect of having three tubes of different lengths to represent the three dimensions of a room was tried. This was exactly equivalent to a room having complete series of axial modes but no non-axial ones, and it was found that the results were very little better than those from a single tube, strongly excited modes being audible up to quite high frequencies.

The conclusion is, clearly, that the higher the frequency of a mode the less likely is it to be individually audible. For a good talks studio, modes above about 300 c/s are seldom distinguishable; in living rooms without properly designed acoustic treat-

ment, the limit is usually higher but in general not more than 1 000 c/s.

The above considerations, applying to the room as a whole, enable one to predict which modes are potentially audible. The two remaining factors listed in Section 2.1 must, however, be considered, since they affect the actual excitation of the modes. The position of the speaker or other source of sound in the room will affect the audibility, since a mode is most vigorously energized by a high-impedance source at a pressure antinode but not by a source at a node. Thus all modes will be equally excited by a source at a corner, but for most other positions there will be many modes that are only weakly excited. The expression 'at a corner' implies here that the source is less than a quarter-wavelength from the corner for sound of the highest frequency, a condition which cannot be realized in practice. A good alternative position, which ensures approximately equal excitation of all modes, is one-third of the way along a diagonal from one corner of the room to the corner farthest away from it. The same remarks apply reciprocally to the position of the microphone used in a broadcasting studio.

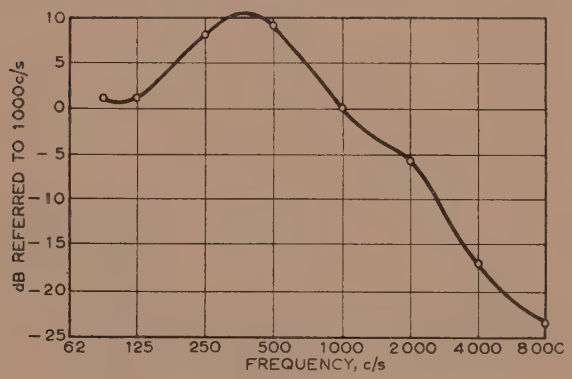


Fig. 2.—Energy distribution in male speech.

Ordinate shows speech power measured in $\frac{1}{3}$ sec intervals (integrated over whole sphere) which is exceeded during 1% of whole time. Derived from data in References 23 and 24.

Lastly we must consider the energy spectrum of speech. The maximum energy, as shown in Fig. 2, is in the neighbourhood of 300–400 c/s. A more important feature, however, is the distribution of the fundamental and formant frequencies, which are considered in Section 3.

(3) COLORATIONS

(3.1) General Characteristics

The characteristic low-frequency colorations which take the form of an unnatural and often monotonous emphasis of certain frequencies in the speaker's voice are to most listeners the most objectionable feature of broadcast speech originating in a small studio, and much effort has been directed towards their elimination by acoustic design and treatment.

In general, a room will have many potential coloration frequencies, determined, as shown above, by the existence of prominent isolated modes. Only a few of these will actually be strongly excited by speech, however, owing to the fact that speech in its lower-frequency region is composed of a limited number of distinct pitches, the voice fundamentals and overtones. These vary from vowel to vowel, with changes in the inflection of the voice and with the individual speaker.

Purely subjective methods of assessing colorations will, therefore, tend to find a smaller number than objective methods, which should, if sufficient representative positions in the studio are examined, show all the modal frequencies of a type capable of giving identifiable colorations, including those of frequencies not strongly excited by the human voice.

(3.2) Subjective Tests for Colorations

The method used by the B.B.C. Research Department is to listen to several people speaking in turn at a microphone in the studio, the voices being reproduced in another room by means of a high-quality loudspeaker. The presence of any over-emphasized tones is noted, and estimates are made of the frequency and severity of each one. A most successful instrumental aid is a selective amplifier which is arranged to amplify a narrow frequency band to a level about 25 dB above the rest of the spectrum. The output is fed in small proportions in parallel with the original signal to the loudspeaker, the proportion being adjusted until it is barely perceptible as a contribution to the whole output. Any colorations can then be heard clearly when the selective amplifier is tuned to the appropriate frequency.

This process is carried out as a routine when testing talks studios, to determine the frequencies of the most obvious colorations. In general, only one or two obvious colorations are found in a studio and it is of interest to see how these are usually distributed.

Fig. 3 shows an analysis of 61 colorations observed in talks

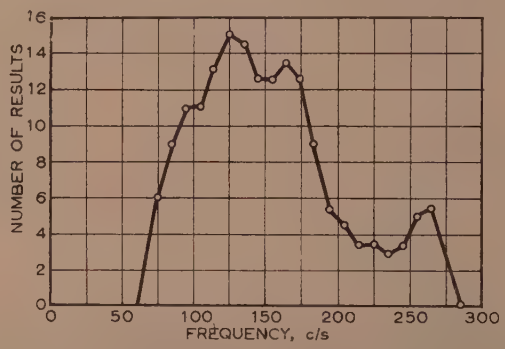


Fig. 3.—Frequency distribution of observed colorations in studios (61 observations).

studios during a period of about two years, using male speech as the programme material. The horizontal scale represents the centre values of frequency bands 10 c/s wide and the vertical scale shows the number of results falling into those bands. To avoid giving undue significance to a chance large number of results in a single column, each ordinate represents the running average with the two on either side of it. It is clear that most colorations fall into the range 100–175 c/s and that there is a subsidiary maximum at about 250 c/s. There is insufficient information on women's voices to plot a similar histogram, but they almost invariably show the strongest colorations between 200 and 300 c/s. Obvious colorations below 80 c/s are very rare, and those above 300 c/s become decreasingly prominent, partly because of the diminution of speech energy and partly because of the increasing numbers of non-axial modes sharing the total energy above this frequency.

The experimental distributions shown in Fig. 3 should be compared with the known spectrum of the voice. Peterson and Barney[3] have measured the fundamental and vowel-formant frequencies for men, women and children. They found that, on the average, men produced fundamental frequencies ranging from 124 to 141 c/s according to the vowel, but that there were large individual variations from this range, the standard deviation being of the order of ±25 c/s. Women, on the other hand, produced average fundamental frequencies ranging from 210 to 235 c/s according to the vowel. The first formants of male

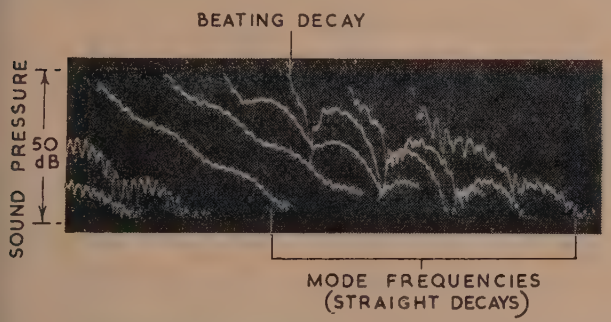
eech start at 270 c/s and those of female speech at 310 c/s, ain with variations of a similar order according to the indi-dual. These figures, compared with the test results of Fig. 3, ggest very strongly that the majority of audible colorations studios are those directly excited by either fundamental or st formant frequencies in the speaker's voice.

(3.3) Instrumental Detection of Potential Colorations

A detailed objective study of the progressive changes in the ecay curve of sound in a studio, as the frequency is varied, as made possible by the development in 1950 by the B.B.C. esearch Department[4] of a logarithmic amplifier with a logarith-ic law over a 60 dB range.

This equipment produces a signal proportional to the logarithm the sound pressure amplitude, which is then displayed against single-sweep time-base on a cathode-ray oscillograph. It ables successive traces with slowly rising tone frequency to be otographed side by side on a moving film, producing con-nuous formations of curves which to a large extent can be terpreted in terms of room modes, mechanical resonances or her features. A description of this method, known as the ulsed glide', with some of the preliminary results, has been ven by Somerville and Gilford.[5]

In the neighbourhood of the frequency of a prominent mode e room behaves exactly as a simple resonant system such as a ned electrical circuit, sound dying away smoothly according to n exponential law which appears on the logarithmic display as straight line. The same behaviour will be observed at an djacent mode, but at frequencies between, both modes will be xcited in opposite phases and beats will appear on the decay rve, which may represent fluctuations of 40 dB or more if the vo modes are excited to a similar extent. An example of the gion between two such modes is shown in Fig. 4. If more



g. 4.—Pulsed-glide displays showing two adjacent modes with beats on intervening decays.

an two modes are simultaneously excited, the beat pattern ll become more complex. Where there are many modes, as a large studio, or a small studio at high frequencies, the ctuations from the exponential law become virtually random. The unfluctuating straight-line displays separated by regions with clearly defined beats are unmistakable indications of strong room modes. Even small rooms show only a small number of such formations, however, out of the hundreds of modes which exist. This verifies the conclusion reached above, that most room modes have little individual effect.

Another characteristic formation is caused by the presence in the room of a mechanically resonant object, such as undamped wall panelling or a metal radiator or lampshade. Such objects are usually forced slowly into oscillation because of high inertia and never become important sources of radiation. However, after the exciting sound has ceased, the vibrations die away more slowly than the room modes, becoming the chief source of sound for the later part of the audible decay. Fig. 5 is an example of a pulsed glide display showing this phenomenon.

To make a clear-cut distinction between these two types of display would be misleading, because wall resonances can give displays similar to those of room modes, and, conversely, an isolated room mode which has very low damping can show a second slope late in the decay curve.

Experience with these methods during the last six or seven years has shown that, provided reasonable precautions are taken to eliminate structural resonances, colorations are almost always associated with room modes. A defect of the pulsed-glide method is that the appearance of the display depends greatly on the position of the microphone, and to assess a room completely it is necessary to repeat the glide at several different microphone positions, deducing the importance of the several features by an inspection of all the displays. A development of the method was therefore tried in which the pressure amplitude display was replaced by one showing the scalar product of the sound pressure and the oscillator signal which produced it. A full account of this test and the results obtained with it were given in a recent B.B.C. Engineering Monograph.[6] The displays, of which Fig. 6 is an example, show characteristic fluctuations, the number of which represents the change of frequency undergone by the sound in its transformation into reverberant energy at a modal frequency. Since the recognizable features of the displays are mainly determined by the frequency information, they are largely independent of the amplitude and consequently are less sensitive to the degree of excitation of a particular mode at each microphone position. It is also probable that the audibility of a coloration is partly connected with the changes of pitch of programme material of neighbouring frequencies, and hence that the test corresponds closely to sensation. It gives rather more reliable predictions of colorations than the simple amplitude display, and is able to resolve neighbouring modes only a few cycles apart. The method has been used more as a laboratory technique than a routine test, which presents difficulties in practice.

(3.4) Design Precautions for Avoidance of Colorations

Having now reviewed the influence of mechanical resonances and of room modes of different types and frequencies, we are in a position to decide on the correct design procedure. Taking

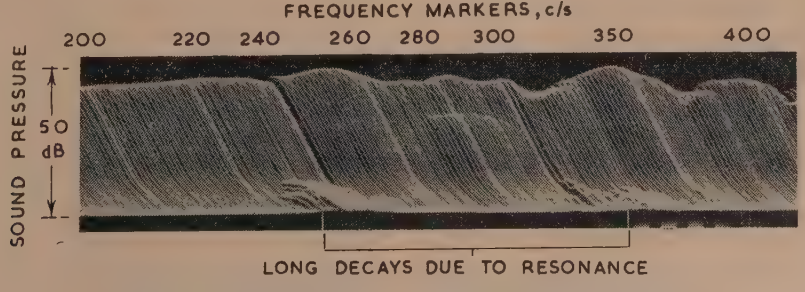


Fig. 5.—Pulsed-glide display showing mechanical resonances.

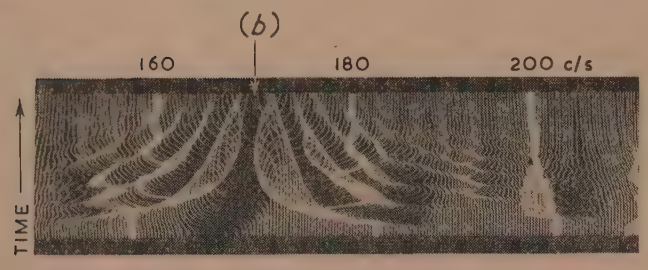


Fig. 6.—Part of coherent glide, showing strong mode at (*b*).

Frequency markers are shown on the top margin, and the horizontal displacement of each trace from zero (approximately at its intersection with the top margin) represents the scalar product of the input tone and output microphone signals.

structural resonances first, little need be said except that structural and lining materials having high Q-factors should be avoided. The materials with which effects of this sort have most frequently been associated in the past are plaster on expanded-metal lathing, plywood, breeze block, and plasterboard with an unsupported back surface.

The problem of room modes is fundamentally more difficult, because they cannot be eliminated or indefinitely reduced without at the same time eliminating all or most of the reverberant sound. There are objections to this, as will be explained below, and we are therefore faced with the more difficult problem of making the unavoidable modes less conspicuous.

Attention must first be given to the dimensions of the room. A simple calculation based on eqn. (1) enables a list of all the axial modes for all three dimensions to be written down in order of frequency. It will be unnecessary to continue the list beyond, say, 350 c/s because, as already noted, the axial modes in a well-designed talks studio will not be prominent above that frequency. The list must next be examined to find modes, or groups of modes with almost the same frequency, which are separated from their nearest neighbours on either side by intervals appreciably larger than their bandwidths. In practice the minimum separation for audibility appears to be about 20 c/s.

Modes or groups separated from their neighbours by greater intervals than this should be noted, and if they fall in the frequency ranges likely to be excited by voice fundamentals and formants attempts should be made to alter the groupings by changes in the proposed room dimensions. It is impossible not to have some isolated groups within the list, but it is usually possible to avoid very bad examples.

Reducing the reverberation time of the room at a particular frequency increases the bandwidth of the modes and reduces their excitation. The application of selective absorption at the frequency of remaining isolated groups would therefore be a possible method of controlling them; this has been tried in the past, sometimes with success, but a limit along these lines is set by the fact that very selective absorbers such as may be required have themselves long decay times and a tendency to reradiate absorbed sound.

However, it is often useful to apply a selective absorber on the walls perpendicular to the longest dimension of the room, because for a given absorption coefficient the axial modes for this dimension will have the smallest bandwidth and will therefore be most likely to be audible.

(3.5) Distribution of Absorption Coefficient between the Boundaries

So far, no account has been taken of the influence of differences in absorption coefficient between one pair of walls and another. In Section 2, it was assumed throughout that all the surfaces had approximately equal average absorption coefficients, and the calculations of the relative importance of the different types of mode were based on this assumption. If, however, all the absorbers are concentrated on two pairs of parallel surfaces, and the third pair is substantially reflecting, strong axial modes will be formed between the latter, all other modes of the room being suppressed. The listener will then hear only one harmonic series of modes, which will remain separate and distinct up to very high frequencies. These conditions give rise to the well-known phenomenon of flutter echo, any time function of the pressure at a source being reproduced periodically with a time interval determined by the distance between the two surfaces in question. It should be emphasized that this can occur only when one pair of surfaces is very much more reflecting than the other two pairs. It is an effect of relative rather than absolute reflection coefficient and does not occur, for example, in a tiled reverberation room, where all the surfaces are highly reflecting. The most familiar example is that of the space between two high walls.

If two pairs of surfaces in a room are highly reflecting and the third pair absorbent, there will be two harmonic series of axial modes, some tangential ones but no oblique ones. The audible modes will extend to considerably higher frequencies than in the case of a uniform room, but the highest-frequency modes will be indistinguishable and there will be no flutter-echo formation. Instead, impulsive sounds will excite a series of clear musical rings up to frequencies of the order of 1 000 c/s.

These conclusions on the behaviour of non-uniform rooms were verified by experiments during the course of the construction of some studios in Portland Place, London. By successive addition of absorbing material it was established that flutters and rings were likely to be noticeable features of the acoustics of any room if the ratio between the mean absorption coefficients of any two pairs of walls was greater than about 1·4 : 1. In designing studios of small or moderate dimensions this ratio should not in any circumstances be exceeded at any frequency, and the aim should be for ratios nearer to unity, with slightly higher mean coefficients for the pair of walls with the greatest separation. The reason for this reservation was given in Section 3.4.

(4) DIFFUSION

Much has been written on the influence of diffusion on the acoustics of studios. A sound field is said to be diffuse if at any moment the intensity of the sound is uniform over the whole volume and if at any point the energy flow is the same in all directions. The implications are that there are no predominant standing-wave systems and that no one position can be distinguished from another. This is an unattainable ideal, but a sound-field may approach it more or less closely, being said to have a greater or less degree of diffusion. The degree of diffusion may be measured by measuring the statistical variation of certain acoustic properties with position, direction, frequency or time. The most satisfactory test, in the author's experience, is to examine the departure of typical decay curves from a generally straight course. Fig. 7 shows two sets of decay curves, (*a*) being from a room with a fairly diffuse sound field and (*b*) from one in which the average damping for modes in one direction was much greater than that for the other directions. In the latter case, the most highly damped modes vanish first, leaving the less damped modes to determine the slope of the later part of the decay.

Many authors have laid down rules for the improvement of studio acoustics by increasing the degree of diffusion. It is usual in many organizations to avoid parallel walls, to introduce irregularities in the wall surfaces and to distribute the absorbing materials as a series of irregular areas over the walls and ceiling. The subjective evidence for the advantages of any of these

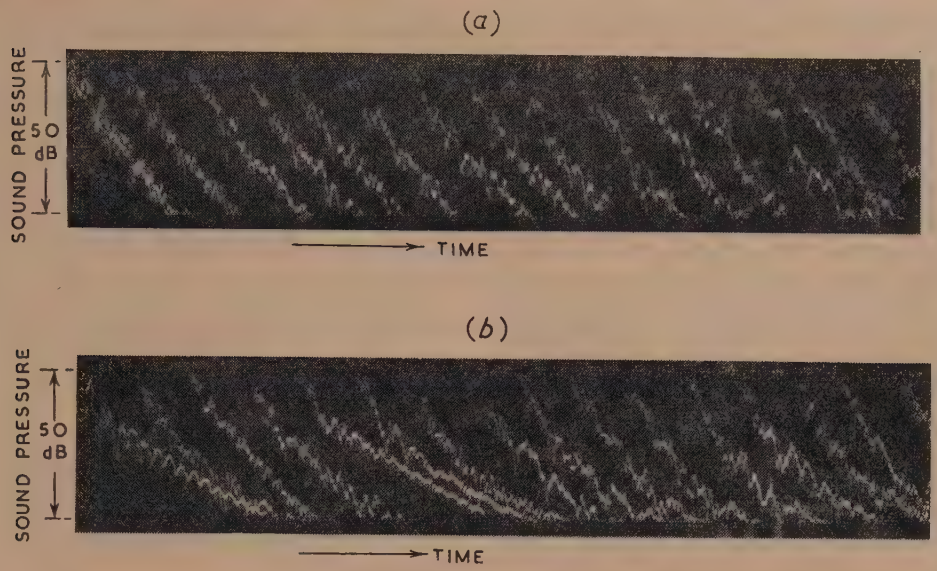


Fig. 7.—Decay curves from diffuse and non-diffuse rooms.
(*a*) Diffuse room: slopes of curves are substantially constant during decays.
(*b*) Non-diffuse room: slopes diminish as decay proceeds.
Each curve represents a different frequency in the range 150–250 c/s.

easures has for the most part been inconclusive, almost all stification having been by reference to theoretical considera-ons or the results of purely instrumental measurements on the coustic properties of the rooms in question.

We will therefore review the evidence in favour of these easures, remembering that the final justification must always e an agreed improvement in quality.

(4.1) Effect of Wall Angle

A common method of increasing diffusion is to build the pposite pairs of walls a few degrees out of parallel. This is sually said to eliminate flutter echoes, but in practice does not o so entirely. It has been reported that a room with non-arallel walls gives a more regular frequency characteristic than rectangular one. Nimura and Shibiyama,[7] working in model ooms of which the shapes could be varied, have measured the requency irregularity[8] which is a measure of the extent to which ne steady-state level in the room due to a source of constant trength fluctuates with frequency. They conclude that non-arallel walls reduce the irregularity below a frequency of about 00 c/s, the maximum effect being produced by an angle of 5° etween opposite walls. Schröder[9] has shown theoretically and xperimentally that the frequency irregularity is directly propor-onal to the reverberation time, and that changes which affect ne reverberation time will therefore also affect the irregularity. low, the absorption coefficient of a surface to grazing-incidence ound is generally half that to normally-incident sound. This as been explained by the author in connection with resonant bsorbers.[10] Non-parallel walls, by discouraging the establish-ent of repeated grazing reflections, will tend to improve the fficiency with which the absorbers act, reducing the reverbera-on time and hence the irregularity.

It must be borne in mind also that the models used for these xperiments had uniform flat walls, whereas this is not normally e case in a studio. It therefore remains to be seen whether the sults are valid under practical conditions.

Full-scale experiments were carried out by the B.B.C. in the rly part of this decade, to determine whether differences of all angle alone were subjectively significant. An experimental studio was built in which the angles of two of the walls could be varied up to a maximum of 6°. Recordings of male and female speech made in parallel and non-parallel configurations were then compared by panels of experienced listeners, but there was found to be no concordant preference for one condition or the other.

It should be remarked here that the wall surfaces were treated by irregular patches of absorbers to give the desired reverberation/frequency characteristic. The conditions were therefore not those of uniform wall surfaces always assumed in the theoretical and experimental evaluation of wall angles. A full-scale experiment avoiding this divergence from 'ideal' conditions is impracticable, since a single uniform absorber with the absorption/frequency characteristic required does not exist.

We will therefore consider now the effects of the perturbation of the individual wall surfaces and the subdivision of absorbing materials as a means of increasing the degree of diffusion in a room, and hence of ameliorating the effects of strong modes on speech quality.

(4.2) Effect of Wall Irregularities

The first work published since the war was that of Somerville and Ward,[11] who showed by means of small-scale models that the perturbation of walls by projections had a measurable effect on the diffusion for sound of wavelengths less than about one-seventh of the height of the projections. Rectangular prisms were found to be more effective than triangular prisms or hemi-cylinders of the same volume, and this result was also demon-strated theoretically by Head.[12]

These conclusions were at variance with the then current opinion that cylindrical diffusers were the most efficient. Meyer and Bohn,[13] for example, found that in free-field conditions hemicylinders produced the greatest scattering effect on plane waves, but free-field results are not necessarily valid in relation to the standing-wave field in a room. The work of Somerville and Ward was later confirmed experimentally by Bruel,[14] and experience in concert halls and large music studios tends to lead to the same conclusions.[15]

However, although in large spaces such as these the breaking up of wall surfaces by projections is both effective and necessary,

it has disadvantages when applied to small studios, since to be effective the projections must be bulky. The wavelength at the commonest coloration frequency (see Fig. 3) is about 2½ m, and to have a worth-while effect the depth of the projections would have to be appreciably greater than one-seventh of this amount, say 50 cm. Not only would such projections be highly inconvenient in a small studio, but the relatively deep recesses formed between them would exhibit their own resonances, a phenomenon often observed with similar accidental features in rooms.

Thus the shallow, irregular, polycylindrical diffusers so often used in studios abroad can hardly be expected to have any effect at all on the most serious bass colorations, though they may have other subjective effects associated with higher frequencies. A few attempts to establish such an effect by subjective tests have been made.

Jeffress, Lane and Seay[16] used two rooms with identical dimensions, one having plane walls and the other polycylindrical. Both were devoid of added absorbing material and both had identical reverberation characteristics. In spite of the difference in wall configuration, however, comparable word-intelligibility tests gave substantially equal mean results of 75·4 and 73·9 respectively.

Within the last few months, a comparison of speech quality between flat and coffered walls has been made in two studios in the B.B.C.'s new centre in Bush House, London. These two studios were designed with the intention of making such a direct comparison possible, and they were therefore arranged to be identical in dimensions, structure, reverberation time and distribution of the absorbing materials. The reverberation characteristics of the two studios are within 0·05 sec of each other at all frequencies. The results of several series of listening tests showed that the audible difference between them was very small—actually less than the variation with microphone position in either studio.

(4.3) Effect of the Irregular Distribution of Absorbing Materials

The other known method of introducing diffusion by scattering is to make the surfaces non-uniform with respect to absorption coefficient, the necessary absorbers being distributed in comparatively small areas. The diffraction effects at the edges of these areas cause the sound to be scattered over a wide angle, thus breaking up the standing-wave patterns. The absorbers are normally much shallower than the diffusers described above, and therefore less wasteful of room spaces. It has been stated, however, that this method is less effective than altering the surface shape,[17] and the B.B.C. Research Department therefore undertook an investigation of the relative effectiveness of the two methods, with full-scale experimental material.

These experiments were conducted in a tiled reverberation room of 27 m^3 volume, and were confined to frequencies above 500 c/s, for which diffusers of reasonable depth would be adequate and room modes not isolated. The criterion for the diffusion was taken as the apparent absorption of an area of absorber entirely covering one wall of the room. With non-scattering walls, the absorption coefficient measured by the reverberation method is low because the decay curves obtained are dominated in the later stages by the modes that avoid incidence at near-normal angles on the absorber.

One wall of a reverberation room was covered entirely with an absorbing material and the total absorption was calculated from measurements of the reverberation time. This was repeated with the same material distributed as five patches on different room surfaces. The latter condition gave a result about 75% higher than when the whole material was on one surface. The whole experiment was repeated, first with twelve stout wooden boxes hung on the walls, and then with the boxes replaced by patches of efficient absorbing material of equal total area. In both cases the addition of the diffusing elements, whether boxes or absorbers, had very little effect on the absorption of the distributed absorber, but increased that of the single-wall arrangement to a figure comparable with that given by the distributed material. The conclusion from these experiments was that, by the criterion adopted, rectangular irregularities and patches of absorber of similar size were equally effective as diffusers within their specific frequency ranges of action. Low-frequency absorbers of the membrane or Helmholtz resonator type are, as a rule, very much shallower than the depth recommended above for projections. They are therefore clearly preferable to projections when good diffusion down to the lowest voice frequencies is required in small studios.

Summarizing the conclusions of this Section, it appears that in small studios no advantage is to be gained from the use of non-parallel walls or diffusing projections, provided that the simpler and more convenient expedient is followed by distributing the absorbers in small areas over as many surfaces as possible.

(5) WHAT MAKES NATURAL SPEECH?

(5.1) The Monaural Chain

Up to the present we have been concerned almost exclusively with the standing-wave systems in small rooms and with the colorations that arise from them. There are, however, other matters of which we must take account in the quest for good speech quality. The basic problem in broadcasting under existing conditions is that there is only one channel between the studio and the listener, whereas we are accustomed to hear speech and music by means of two ears which together supply information about the direction and position of the source.

When listening directly with two ears one is provided with an automatic mechanism for partially rejecting sound other than that coming from the direction of the source to which one is listening. This is an evolutionary faculty possessed by all the higher animals; when it is inhibited by having only one channel of information one is conscious of the reverberant sound and extraneous noise to such an extent that speech loses its intelligibility and music its definition unless steps are taken to increase the ratio of direct to reverberant sound.

(5.2) Reverberation Time

The subjective balance between the direct sound and unwanted sound made up of reverberation and extraneous noise may be restored most easily by reducing both these components. The reduction of extraneous noise is a matter of providing effective sound insulation for the studio and giving attention to the noise generated by ventilation systems and other sources inside the room.

The reverberant sound is reduced simply by the application of sound absorbers. A certain amount of low-frequency absorption is provided by the compliance of the room structure itself, sound energy being dissipated by frictional losses as the walls, floor and ceiling vibrate. Unfortunately, good sound insulation in practice necessitates building massive walls of low compliance, and in these circumstances there is no gratuitous absorption, such as is derived from the floors, ceilings and windows of ordinary houses.

It is therefore necessary to add considerable extra absorption to reduce the reverberation time to a point where speech by the monaural listening chain has the intimacy of conversation in a well-furnished room.

One important effect of reverberation should be noted here. the studio is very non-reverberant or 'dead', the resulting peech comes from the listener's loudspeaker substantially nchanged by any added reflections from the studio walls. It then modified by the reverberation in the listener's own room exactly the same way as the voice of an occupant, and the lusion of actual presence is created. Conversely, reverberation dded in the studio will be recognizable as foreign to the room nd will give the illusion that the loudspeaker is a hole in the all communicating with another room where the broadcaster sitting. Different individuals have definite preferences for one r other of these conditions, and, in the author's own experience, ost of those engaged in the engineering or production aspects f broadcasting in this country prefer the illusion of presence. should be remarked, however, that Continental broadcasting rganizations seem to prefer a very much more reverberant ound.

Whether the individual's preference could be correlated with is general psychological make-up is an interesting speculation utside the competence of an engineer. The broadcaster himelf will, however, disagree with the implications of the majority noice, because to speak in a very dead studio is unpleasant, apping the confidence of all but the most experienced news eader and compelling the speaker to raise his voice in an effort o obtain the reassurance given by the reflections which would einforce the voice in other circumstances.

The type of microphone is important in this connection. An mnidirectional microphone, such as a moving-coil or piezolectric instrument, increases the ratio of reverberant to direct ound, thus requiring a deader studio and increasing the strain n the broadcaster. Ribbon microphones with a figure-of-eight haracteristic are rather better, since they reject part of the everberant sound, and the directional characteristics may be sed to discriminate against the strongest modes. Normally, here the diffusion in the studio is fairly good, a ribbon microhone placed diagonally across the room is usually found to be est since it reduces axial modes from the two horizontal direcons by about 3 dB with respect to the rest, whilst discriminating lso against up-and-down modes. A cardioid microphone educes the apparent liveness of the studio but does not much lter the relative effects of different modes.

(5.3) Shape of the Reverberation Characteristic

The quality of speech is critically dependent on the shape of ne reverberation-time/frequency characteristic. Past experience nd controlled experiments have combined to show that the everberation time should be independent of frequency from 0 c/s to 8 kc/s. Deviations from a level characteristic are easy o recognize; excess of low-frequency reverberation produces boominess' and distinct colorations, while excess in the region 50–500 c/s is most unpleasant, giving a throaty, strangled uality to speech. A long reverberation time (above 2000 c/s) roduces sibilance or 'breathiness'. A slightly drooping characeristic below 500 c/s is perhaps ideal, giving the most natural peech quality.

(5.4) Microphone Correction Circuits

It is a common device to include in the microphone circuit a ter to give slight bass attenuation. This is usually carried out trial until acceptable speech, free from boominess, is obtained. orrection may be required in some cases to compensate for a se in the reverberation time below, say, 100 c/s, but there is nother reason. Most people listen to broadcast speech at a gher level than that of the broadcaster's voice in the studio or a friend speaking to them in person. Somerville and ownless[18] found that members of the public preferred to listen to an average level of 71 dB above the standard reference level of 10^{-16} watt/cm^2, whereas the typical level of conversational speech is about 60–65 dB. A well-known property of the ear[19, 20] is that the equal-loudness contours plotted against frequency approach each other closely at low-frequencies. Speech will therefore sound bass-heavy if it is reproduced at an unnaturally high level. A bass cut of 3 dB at 50 c/s relative to 250 c/s would be of the right order to correct for the difference between average listening levels for broadcast speech and live conversation. A pressure-gradient microphone closer than about 60 cm also requires a cut of a similar amount to correct for the curvature of the wavefronts. Compensation for long reverberation in the bass by further electrical equalization is possible to a limited extent only, since the effects of colorations, being frequency-selective, can only be removed by cuts of such a magnitude that an emasculated quality is imparted to the speech.

(5.5) Influence of the Listening Room

The listener's own room has an influence on the transmitted speech since it adds colorations and other effects of reverberation in the same way as the studio. However, although any additions from this cause are reduced by the binaural rejection mechanism and assume relatively less importance, they are an impediment to critical listening and will therefore be considered in Section 7.3.

(6) THE DESIGN OF TALKS STUDIOS

(6.1) Size and Shape of Studios

The dimensions of a talks studio should be large enough to give reasonably close spacing to the axial modes. Volumes from 1500 to 4000 ft^3 (43–114 m^3) are generally satisfactory, bad colorations being difficult to avoid in studios below this range, and larger studios giving insufficient advantage to justify the increased expense of construction and treatment.

It is doubtful whether any of the preferred ratios between dimensions, published from time to time, can be upheld. Certainly the once popular 5 : 3 : 2 ratios will normally give at least one isolated group within the worst frequency range. Neither can one depend on proposed ratios based on a consideration of modal frequency-spacing statistics,[21] since the presence of a single isolated group, which may not greatly affect the mean frequency spacing, will give a serious coloration. The only way that has had any success, in the author's experience, is to work out the axial-mode frequencies for a set of trial dimensions, as described in Section 3.4, and to adjust the dimensions until a satisfactory mode spacing is achieved. Tests of these several theories were made in an experimental studio in the B.B.C. Research Department in which one of the walls, constructed of heavily reinforced clinker block, could be moved perpendicularly to its plane to give dimension ratios predicted by various theories as 'good' or 'bad'. The results confirmed substantially the overriding importance of isolated axial modes, but failed to show any difference between configurations with high or low values of the frequency-spacing statistic.

The shape of the studio may be rectangular, or the walls may be splayed at angles up to a few degrees if the shape of the site renders this more convenient. There is insufficient evidence for a dogmatic view about these questions, and one cannot rule out completely non-rectangular shapes such as triangular or pentagonal prisms, which have been used by other organizations. There have been suggestions, however, that by their unfamiliarity in ordinary homes, such shapes impart a somewhat unnatural quality to speech. Irregularities in the wall shapes are, as shown in Section 4.4, unnecessary.

(6.2) Reverberation Time

The reasons for the choice of reverberation time have been dealt with in Section 5.2. B.B.C. experience has shown a time of about 0·3 sec to be the optimum, with a possible slight reduction below 300 c/s. The additional bass absorption required to give this reduction is difficult to achieve without making full use of structural absorption from the walls, floor and ceiling of the studio. Extra absorption is obtained by the use of resonant membrane absorbers consisting of sheets of bituminous roofing felt sealing an air space.[10] Helmholtz resonator absorbers have also been used, since they have the advantage that they can be tuned to the required frequency and adjusted for bandwidth and maximum absorption by the addition of resistive materials in the necks.[22]

Continental talks studios, as pointed out also in Section 5.2, tend to be more reverberant, and times up to 0·5 sec or even more are encountered.

(6.3) Application of Absorbing Materials

The aim, as described in Section 3.5, is to distribute the low-, middle- and high-frequency absorbers equally on each of the pairs of parallel wall surfaces, except that low-frequency absorbers should be slightly in excess on the end walls, i.e. those separated by the longest distance. On the individual surfaces there should be patches of efficient absorber surrounded by areas of poor absorber such as hard plaster. The absorbers should be arranged on the individual pairs in such a manner that there are no large areas of reflecting surface directly facing each other, since these will give rise to flutters.

The effectiveness of any absorber depends markedly on the position it occupies in the room. Calculation of the reverberation characteristic of a small room is therefore never very precise, and provision must be made for adjustment of the absorbers after completion of the studio. It is the usual practice in the B.B.C. to construct all absorbers, wherever possible, with detachable covers which can be removed for alteration, replacement or removal of the contents. The cover itself may act as a filter or the inductive element of a Helmholtz resonator, in which case it is made of a perforated board. Alternatively, it may be of fabric. Perforated covers tend to limit the performance of the absorber at high frequencies, thereby producing excessive sibilance, while fabrics soon become dirty and may shrink while being cleaned so that they cannot be refitted. This disadvantage has been overcome in a recent B.B.C. design in which the fabric is stretched over a detachable frame arranged to accommodate surplus fabric to compensate for shrinkage.

(7) DESIGN OF CONTROL CUBICLES AND OTHER LISTENING ROOMS

(7.1) Design Principles

It has already been said that, from the point of view of acoustic faults and the methods of overcoming them, listening rooms may be treated in the same way as small studios. Their different function, however, imposes differences in design, which will be briefly considered in this Section.

(7.2) Resemblance to Living-Room Conditions

Unlike electronic and electro-acoustic equipment, for which one can set an ideal input/output characteristic as a goal for both broadcasting organizations and their listeners, a domestic living room must be accepted very much as it exists. Control-cubicle and quality-monitoring-room acoustics should therefore not be very dissimilar from the average conditions encountered in private houses. A few years ago an acoustic survey of living rooms was carried out, embracing various types of construction from Georgian houses in Pimlico to post-war maisonettes with solid floors in the suburbs. Fig. 8 shows the mean reverbera-

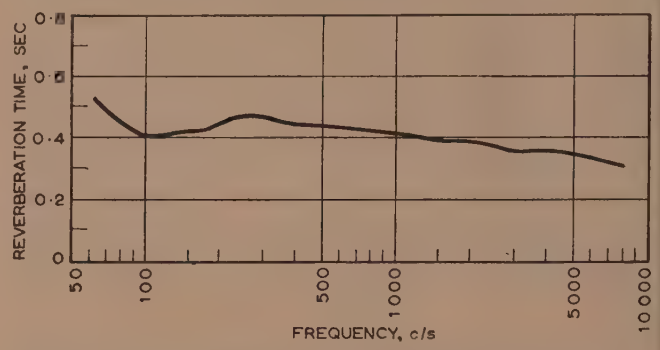


Fig. 8.—Mean reverberation characteristic of 16 living-rooms.

tion characteristic for all the rooms surveyed. There was surprisingly little variation in the results for conventionally furnished rooms, though the reverberation times of rooms with joist floors averaged about 0·05 sec less than those of the immediate post-war houses with solid floors built directly on to concrete foundations. Listening tests with recordings of speech from several studios played into the rooms showed that most of them allowed the characteristics of speech from the different studios to be distinguished, but some of the rooms introduced very severe colorations which entirely masked other effects. It was decided to adopt the mean curve of Fig. 8 as a pattern for listening rooms, pending the completion of subjective tests on the effects of listening-room acoustics on the judgment of speech quality.

(7.3) The Effects of Listening-Room Acoustics on Judgment of Quality

The fact that the listening room does not have a predominant effect on quality is very largely due to the binaural mechanism. This is particularly true of rooms used for listening to music programmes, since, objectively speaking, the standing-wave effects in the room can accentuate particular notes by as much as 8 or 10 dB relative to their neighbours. Fig. 9 shows two pulsed glides obtained by radiating the test tone into a studio, transmitting it by the studio microphone and the listening-room loudspeaker and photographing the resulting decay curves as received by a microphone in the listener's position. The two glides were obtained in the same listening room, but Fig. 9(*a*) was from an anechoic sound-measurement room and Fig. 9(*b*) was from a normally treated talks studio. It will be seen that the glides are sufficiently similar to suggest that it was the listening room which had the principal effect on the overall acoustics.

Accordingly, since it was common experience, verified by the living-room tests described above, that studio differences were nevertheless distinguishable, subjective tests were undertaken to determine to what extent the binaural mechanism enabled one to reject the cubicle acoustics.

Recordings from six studios were played in four different listening rooms to panels of engineers. The test programme consisted of short passages read partly from one of the six studios and partly from another, and the subjects were required to state a preference for one part of the passage or the other. Each studio was compared with every other in this way, making fifteen paired comparisons in all. Statistical analysis of the answers revealed the following information:

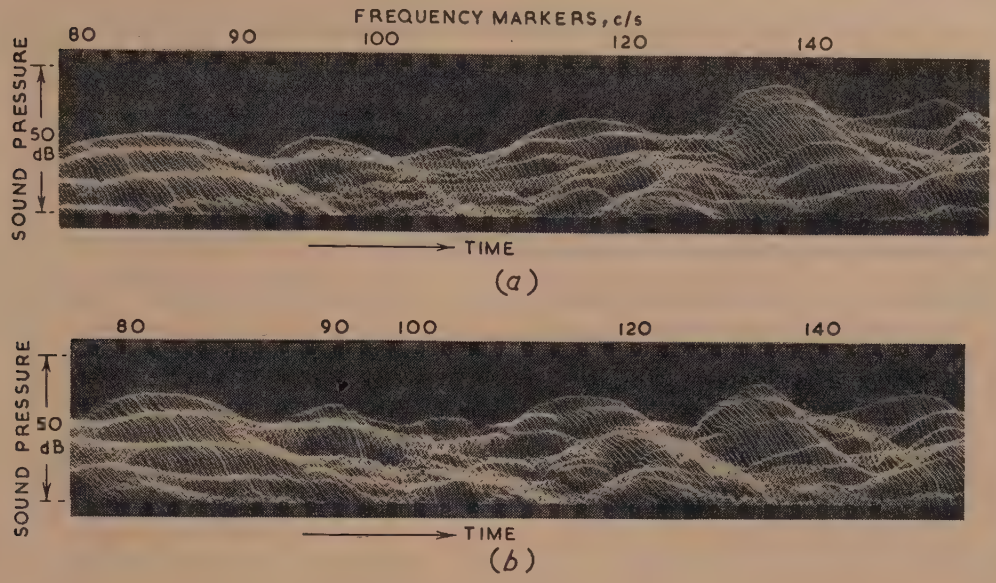


Fig. 9.—Combined pulsed glide of studio and listening room.

(*a*) Normal talks studio and listening room.
(*b*) Non-reverberant room and listening room.

(*a*) The average order of preference of the six studios.
(*b*) Differences in the rank order produced by using different stening rooms.
(*c*) The self-consistency of the individual subjects.
(*d*) The effect of the listening room on self-consistency.

The results showed that an over-reverberant listening room, hile not altering the order of average preference, greatly creased the inconsistency of the subject's answers. A room ith a long reverberation time in the bass favoured a studio with heavy bass cut, while a very dead room favoured a studio with ng bass reverberation, presumably because a listener sitting pproximately on the axis of the loudspeaker (the usual position r critical listening) received an excess of high-frequency sound.

The longest reverberation time which can be permitted without n adverse effect on either the order of preference or the consstency of judgment was approximately 0·4 sec, and it is to be oted that the average curve of Fig. 8 does not rise far above is value.

As the result of these investigations, all control and listening ooms are now designed to have reverberation times of 0·4 sec o to 1 000 c/s, falling steadily above this frequency to 0·3 sec 8 000 c/s.

It goes without saying that the same care must be exercised avoid serious colorations in listening rooms as is necessary the case of studios.

(8) CONCLUSIONS

The acoustic design criteria for small studios and listening ooms are given in the last two Sections of the paper. The imary requirement is to avoid the occurrence of axial modes parated by frequency intervals large compared with their ndwidths. Such isolated modes are generally audible as olorations if they occur in the neighbourhood of the fundaental and formant frequencies of speech. It is impossible to iminate them entirely from small rooms without reducing verberation to an undesirable extent, but they can be neliorated by correct design.

Flutters and rings at higher frequencies are usually due to marked differences in the absorption coefficients between one pair of opposite boundary surfaces and the other two pairs.

Good diffusion of the sound field is accepted as necessary, and is best achieved by arranging the absorbing materials in irregular patches on the individual surfaces. The addition of diffusers as commonly understood is then unnecessary.

In listening rooms and control cubicles, the binaural listening conditions allow a longer reverberation time than in studios, though excessive reverberation makes critical listening more difficult.

(9) ACKNOWLEDGMENTS

The author wishes to thank the Director of Engineering of the British Broadcasting Corporation for permission to publish the paper. He also acknowledges the work of his colleagues in the Research Department, which has contributed much of the material of the paper.

(10) REFERENCES

(1) MAYO, C. G.: 'Standing Wave Patterns in Studio Acoustics', *Acustica*, 1952, **2**, p. 49.

(2) AXON, P. E., GILFORD, C. L. S., and SHORTER, D. E. L.: 'Artificial Reverberation', *Proceedings I.E.E.*, Paper No. 1796 R, September, 1955 (**102** B, p. 62).

(3) PETERSON, G. E., and BARNEY, H. L.: 'Control Methods used in a Study of the Vowels', *Journal of the Acoustical Society of America*, 1952, **24**, p. 175.

(4) BEADLE, D. G., and MAYO, C. G.: 'Equipment for Acoustic Measurements, Part 4, The Direct Measurement of Reverberation Time', *Electronic Engineering*, 1951, **23**, p. 462.

(5) SOMERVILLE, T., and GILFORD, C. L. S.: 'Cathode Ray Displays of Acoustic Phenomena and their Interpretation', *B.B.C. Quarterly*, 1952, **8**, p. 1.

(6) GILFORD, C. L. S., and GREENWAY, M. W.: 'The Application of Phase-Coherent Detection and Correlation Methods to Room Acoustics', B.B.C. Engineering Monograph No. 9, 1956.

(7) NIMURA, T., and SHIBAYAMA, K.: 'Effect of the Splayed Walls of a Room on the Sound Transmission Characteristics', *Journal of the Acoustical Society of America*, 1957, **29**, p. 85.
(8) BOLT, R. H., and ROOP, R. W.: 'Frequency Response Fluctuations in Rooms', *ibid.*, 1950, **22**, p. 284.
(9) SCHRÖDER, M.: 'Die Statistischen Parameter der Frequenzkurven von Grossen Raümen', *Acustica*, 1954, **4**, p. 594.
(10) GILFORD, C. L. S.: 'Membrane Sound Absorbers and their Application to Broadcasting Studios', *B.B.C. Quarterly*, 1952–53, **7**, p. 246.
(11) SOMERVILLE, T., and WARD, F. L.: 'Investigation of Sound Diffusion in Rooms by Means of a Model', *Acustica*, 1951, **1**, p. 40.
(12) HEAD, J. W.: 'The Effect of Wall Shape on the Decay of Sound in an Enclosure', *ibid.*, 1953, **3**, p. 174.
(13) MEYER, E., and BOHN, L.: 'Schallreflexion an Flächen mit Periodischer Struktur', *ibid.* (Beiheft), 1954, **4**, p. 195.
(14) BRÜEL, P. V.: 'Automatic Recording of Frequency Irregularity in Rooms', *ibid.*, 1954, **4**, p. 21.
(15) SOMERVILLE, T., and GILFORD, C. L. S.: 'Acoustics of Large Orchestral Studios and Concert Halls', *Proceedings I.E.E.*, Paper No. 2297 R, March, 1957 (**104** B, p. 85).
(16) JEFFRESS, L. A., LANE, R. N., and SEAY, F.: 'Effect of Wall Irregularity on Intelligibility', *Journal of the Acoustical Society of America*, 1955, **27**, p. 787.
(17) MORSE, P. M., and BOLT, R. H.: 'Sound Waves in Rooms', *Reviews of Modern Physics*, 1944, **16**, p. 137.
(18) SOMERVILLE, T., and BROWNLESS, H. J.: 'Listeners' Sound Level Preferences', *B.B.C. Quarterly*, 1949, **3**, p. 12.
(19) FLETCHER, HARVEY: 'Speech and Hearing' (Macmillan, London, 1929), p. 225 *et seq.*
(20) MINTZ, F., and TYZZER, F. G.: 'A Loudness Chart for Octave Band Data on Complex Sounds', *Journal of the Acoustical Society of America*, 1952, **24**, p. 80.
(21) BOLT, R. H.: 'Normal Frequencies in Rectangular Rooms', *ibid.*, 1946, **18**, p. 130.
(22) WARD, F. L.: 'Helmholtz Resonators as Acoustic Treatment in the New Swansea Studios', *B.B.C. Quarterly*, 1952, **7**, p. 7.
(23) DUNN, H. K., and WHITE, S. D.: 'Statistical Measurements on Conversational Speech', *Journal of the Acoustical Society of America*, 1940, **11**, p. 278.
(24) DUNN, H. K., and FARNSWORTH, D. W.: 'Exploration of Pressure Field around the Human Head during Speech', *ibid.*, 1939, **10**, p. 184.

DISCUSSION BEFORE THE RADIO AND TELECOMMUNICATION SECTION, 15TH DECEMBER, 1958

Mr. J. Moir: It is a major difficulty with problems of room acoustics that the final judgment of the room performance is almost entirely subjective. We know a great deal about the objective performance of a room, but we do not know what the objective performance should be in order to give an excellent result when subjectively judged.

Most of the paper is concerned with the acoustic design of talks studios and listening rooms. It might be qualified by saying 'listening rooms for talks', for there is very little comment about the conditions required for the reproduction of music.

Room modes are undoubtedly the main troubles in a small room, and the paper is particularly valuable in indicating that the axial modes represent the major difficulty.

I doubt whether the reverberation time at low frequency has much meaning. When we measure the reverberation time of a room, particularly in the region below about 150c/s (depending on the room size), the results generally vary radically with changes of frequency of only a few cycles per second. Curves commonly presented show a smoothed shape for the reverberation-time/frequency relation, but they are not so smooth in practice. It is not unusual to find that the reverberation time changes by a factor of 2 : 1 over a frequency range of a few cycles per second, and it is often doubtful whether the high or the low value is the effective one.

There has been a radical change of thought in recent years over the best shape of reverberation-time/frequency curve. At one time the optimum was assumed to be something rising fairly sharply at the low-frequency end of the range. In recent years, particularly as a result of experience with the Royal Festival Hall, there has been a tendency to favour a curve which falls off a little at the low-frequency end.

It has been claimed—and I think that measurements support it—that the majority of the old concert halls had reverberation-time/frequency curves which rose sharply at low frequency. I do not like music rooms in which the reverberation time falls off at low frequency, for the music appears to lack 'body' and roundness. There is no doubt that for the reproduction of speech, where intelligibility is the main factor, a reverberation-time/frequency curve which falls off at low frequency always gives better results than one which rises at low frequency; but music is often just as important as speech, and one may have to compromise.

Flutter echoes have also been mentioned. They are very prevalent in cinemas. A hand clap near one wall is followed by a series of 20 or more 'slaps', and one is left with the impression that the acoustic conditions are poor, but we have never found any real trouble from these flutters. I would like to have the author's comments on how troublesome they are in practice, and indeed if they are a serious problem in studios.

Ringing at high frequency is much more serious. There are very few living rooms in which there is not some evidence of this. In my experience, non-parallel walls in small or large rooms seem to produce a greater reduction in flutter echo and ringing than the author seems to have found. Non-parallelism of 5–10° appears to remove most of the flutter echo and ringing.

Finally, in my view, if a room requires extensive treatment for stereophonic listening there is something wrong with the stereophonic equipment or the recording. The better the stereophonic reproduction system, the less trouble we have with room acoustics.

Mr. H. R. Humphreys: I am rather concerned by the general inference that there is *one* good design for a talks studio. The reproduction of speech must be regarded as an artifice, since, however natural the speaker may sound, he is not, in fact, there. I think that this is justification for some variation in studio acoustics.

The author suggests a mean reverberation time of 0·3 sec, but is this the right policy? Let us consider the different uses to which talks studios are put. There is the single speaker, perhaps an announcer, and one might regard this as unaccompanied voice. There is discussion—duet, trio, quartet, etc. Then there is the commentator, which one might regard as a concerto for voice and accompaniment.

A single speaker's voice may perhaps be allowed to issue at the receiving end as 'disembodied' speech, but if this is done for a group of speakers in discussion we lose all sense of perspective. Only by using a limited and carefully controlled amount of studio acoustics can one get some sense of perspective on a monaural channel.

The commentator will most likely be accompanied from time

time by some 'noises off' or will be dubbed over a musical sound-effects background which will itself set the acoustical cene. The major requirement for the voice is then intelligibility, hich can probably best be achieved against the distraction of ackground by using a fairly 'dead' or completely 'dead' studio.

With regard to 'loud random reflections', a few weeks ago r. Beranek spoke on concert-hall acoustics, and he made a oint about first reflections in these rooms. Does the author elieve that the first reflections in talks studios have any impor- ance, bearing in mind that the time intervals will be very much horter because of their comparatively small size?

The author states that, in some cases, structural resonances re caused by certain surfaces having a longer reverberation time han the room itself, and he mentions 'plasterboard with an nsupported back surface'. I do not know what he means by his because I cannot imagine using plasterboard without upporting its back surface in some way or other.

In rooms with non-parallel walls there is a tendency for the ffectiveness of absorbents to be slightly increased. How much xtra absorption should one allow in designing for reverberation ime when the walls are non-parallel?

Mr. C. G. Mayo: A studio, like a violin or loudspeaker, can nly respond in its own eigentones. Rather than being critical, am therefore surprised how well the studio reproduces the ounds emitted. If a note is sounded at a frequency between hose of two eigentones reasonably close together, the response vill have the pitch of the note sounded, and not that of either igentone. The pitch of the response is attained by a kind of eat frequency having a rapidly changing envelope. The zero rossings give the right pitch, but the frequency is not that ssociated with the zero crossings.

A room having a very large number of eigentones or modes an likewise respond to any frequency. If the modes were qually spaced in frequency (as in an open-circuited cable) the ound emitted would be precisely reproduced, but with a time elay or echo. When eigentones are closely and randomly paced, however, the input is faithfully reproduced. Thus a iolin or loudspeaker gives a remarkably good response in spite f its limited power.

In remote ages, the ear was used primarily for the rough ocation of position and the determination of the size of an nclosure like a cave. The ear instantly estimates critically the ize of a talks studio or an auditorium.

It is the absence of modes rather than their presence that spoils small talks studio. A prominent single mode is annoying ecause it lacks the necessary help of other modes.

Mr. P. P. Eckersley: Why do so many people turn their oudspeakers up louder than normal speech level? It may be hat the top frequencies of loudspeaker reproduction are usually ttenuated far more than in direct speech, but this suggestion oes not look particularly viable in view of the fact that most eople turn their tone controls to 'mellow'.

It is sad to realize that the B.B.C. takes all this trouble to get he acoustics of the studio to near perfection while the listener ullifies all its efforts.

Mr. W. West: For testing loudspeakers, I use the criterion that a source of sound (e.g. a talker), such as could be present in the listening room, should reproduce as though it were in the room. With this criterion any observable effects of the talks studio are unwanted, and thus ideally the studio should be non-reverberant. Recordings of speech that we have made in a non-reverberant room are better, in respect of clarity and signal/noise ratio, than those made in normal rooms acoustically treated as studios. The objection that a speaker does not like talking in a non-reverberant room is unreal; the room has a calm, relaxed atmosphere, and the idea that there is no need to talk loudly is easy to put across. If the talker likes to wear a headphone to hear his voice by sidetone, the loudness of his talking can be controlled, without his knowledge, by adjusting the amount of sidetone he receives.

In Section 5.1 it is stated that binaural hearing reduces the effects of reverberant sound in the listening room. What is called, for brevity, the binaural mechanism is, of course, more than a mechanism, because subtle faculties of the brain are involved. Has the author any experience to indicate whether a so-called stereophonic system, using loudspeakers in the listening room, is of any help in rejecting reverberant sound coming from the studio?

Mr. P. F. Cook: With regard to the intensity of reproduction of a voice from a loudspeaker I suggest that this is largely a psychological matter. In normal conversation our speech content is full of redundancies and the information content is low, but it is very noticeable that, if someone in a group of people conversing in a room says something interesting, the others immediately look at him. This illustrates the fact that we get a lot of information from a speaker present in a room by watching his lip movements and general expressions. When that visible information is removed we feel the need of increased intelligibility. The normal reaction, since the broadcast message usually has a higher information content than ordinary conversation, is to turn up the volume. As soon as the broadcast speaker becomes uninteresting and the people in the room want to discuss things among themselves, the volume tends to be turned down.

I suggest that our normal conversation is conducted at an intensity which is on the borderline of intelligibility and that omissions are made good by the visible factors. When these are absent we choose the increased intelligibility—particularly that associated with the higher frequencies—which result from increased volume.

Major W. V. G. Fuge: It is well known that, if we remove the furniture and carpet from quite a small room, speech has an echo added to it. The less furniture there is, the greater the echo effect. Has any use been made of this phenomenon for adding a pleasing amount of echo to speech, by emitting it from a loudspeaker at one end of such a room and picking it up by a microphone at the other end with the echo added. Can such a desirable added effect be made to predominate over the acoustic effects due to the room in which the speech is made, and to the room in which it is heard?

THE AUTHOR'S REPLY TO THE ABOVE DISCUSSION

Mr. C. L. S. Gilford (*in reply*): I agree with Mr. Moir that the nal judgment of a room must be subjective; all the objective ork described in the paper was carried out with continual eference to parallel subjective experiments. Most of the ıbjective work on listening rooms was carried out on speech, nce only speech is normally broadcast from studios having wer reverberation times than ordinary living rooms. The coustics of the living room would be expected *a priori* to have less importance when listening to programmes emanating from much more reverberant studios, and this is borne out in practice. Measurements of reverberation time at low frequencies have more meaning than he suggests, because repeatable results can be obtained by averaging readings from several microphone positions using a test sound of finite bandwidth and also measuring the decay time of particular modes. I agree that, in practice, flutter presents much less trouble than

ringing, but it is not my experience that non-parallel walls reduce ringing.

Mr. Humphreys is right to point out that no studio can be equally correct for all kinds of speech. The requirement which has been borne in mind is that of didactic speech, news bulletins, etc., which comprise the majority of speech programmes. Microphone placing can, to some extent, introduce the perspective required for discussions or the intimacy of a narration. Narrators' studios associated with drama suites are designed to be less reverberant than ordinary talks studios.

With regard to early reflections, there is no doubt that those from a hard table top or a cubicle window too close to the microphone alter speech quality, but I am not convinced of their importance in determining the characteristics of a studio as a whole. There is room for further investigation. An isolated complaint of coloration was actually traced to unsupported plasterboard, and I am glad to know that it is unlikely to happen again.

In reply to Mr. Eckersley, I have found that a subject who is asked to adjust a loudspeaker, so that a colleague's voice reproduced by it appears of the correct loudness, will invariably set the level too high—sometimes by as much as 8–10 dB. I have always imagined this to result from psychological influences such as increased attention, in the case of direct speech, due to the physical presence of the speaker. I therefore agree entirely with Mr. Cook's most informative remarks.

Mr. West's statement about the effects of acoustics on the speaker conflicts with my experience. The majority of regular and casual broadcasters find a very dead studio unpleasant, and although an inexperienced broadcaster can easily be trained to speak quietly, it is better that conditions should be acceptable from the start.

Some experiments with a crude stereophonic system carried out some years ago were inconclusive. It would be worth while to repeat them using better studios and a modern stereophonic system.

Major Fuge's idea of using a supplementary room to add reverberation is, in fact, in everyday use in broadcasting and recording organizations. The room is usually walled with smooth concrete or tiles to give it a long reverberation time, but some absorption may be introduced to adjust the frequency characteristic.

21.396.946:621.396.8

The Institution of Electrical Engineers
Paper No. 2963 E
May 1959

OBSERVATIONS ON THE U.S.S.R. EARTH SATELLITES AND THE STUDY OF RADIO-WAVE PROPAGATION

By W. C. BAIN, M.A., B.Sc., Ph.D., and E. D. R. SHEARMAN, B.Sc.(Eng.), Associate Members.

(*The paper was first received 8th November, 1958, and in revised form 27th February, 1959.*)

SUMMARY

Measurements of bearing, angle of elevation and Doppler frequency hift have been performed at Slough on transmissions from the ussian earth satellites. The examples considered here indicate that he observed phenomena can be accounted for satisfactorily in terms f our present knowledge of ionospheric propagation. A short section included on orbital determination from the measurements.

(1) INTRODUCTION

At the Radio Research Station, Slough, a number of observaons have been made on radio transmissions from the Russian arth satellites. Most of these have been carried out in order o increase our knowledge of radio-wave propagation in the onosphere, but a certain amount of work has also been directed o the provision of some results early in the life of a satellite, o that a first approximation to the orbital parameters might be btained quickly. The measurements described here are of earing and angle of elevation at 20 Mc/s, and of the received requency at both 20 and 40 Mc/s. Three examples will be iven of the investigations carried out.

(2) EQUIPMENT

Bearing at 20 Mc/s.—The bearing observations at 20 Mc/s ere carried out with a standard h.f. Adcock direction-finder. he receiver used has twin channels and a cathode-ray-tube isplay. The reading of bearings was done visually and was med by a chronometer.

Angle of Elevation at 20 Mc/s.—Two horizontal loops spaced ertically were connected to the channels of a twin-channel eceiver. The angle of elevation could then be derived from a easurement of the amplitude ratio of the signals at the two erials. There may be up to four ambiguities in the interpretaon of the ratios, but it was found possible to eliminate these by omparing the results with other measurements. Further details f the equipment are given by Wilkins and Minnis.[1]

Frequency at 20 and 40 Mc/s.—The frequency of the transmisons was measured by applying the aerial outputs to the vo receivers, together with standard reference frequencies of 0·003 and 40·000 Mc/s respectively. These reference freuencies were generated from a high-stability (1 part in 10^8) 00 kc/s signal by means of a frequency synthesizer. A beat ote of 2 kc/s ± the Doppler frequency was obtained from the udio-frequency output of the receiver and was recorded on agnetic tape, together with timing signals. The Doppler equency was subsequently measured by playing back the tape cord and matching the audio frequency from a calibrated enerator to the signal with the aid of a cathode-ray tube disaying Lissajous figures. This analysis procedure was repeated veral times to reduce the errors in measurement.

(3) RECEPTION OF SIGNALS FROM A DISTANT SATELLITE

When the radiated power is sufficiently great, the transmissions from a satellite can be used to study propagation over great distances. One example of the reception of signals when a satellite was far from Slough occurred on the 7th October, 1957, near 1130 U.T. The Adcock direction-finder was operating at 20 Mc/s at the time, and a bearing was first obtained on the transmissions from the satellite 1957α at 1129 U.T., as shown in Fig. 1. Although the signal never became very strong, bearings

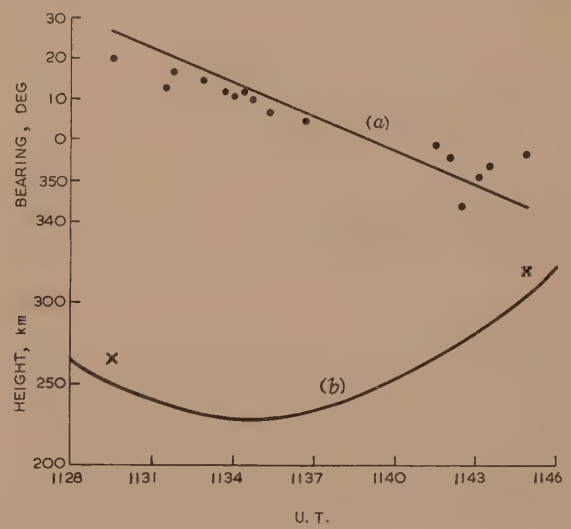


Fig. 1.—Bearings and heights for the first Russian satellite 1957α on 7th October, 1957.

● Observed on 20 Mc/s.
× Estimated true heights of maximum ionization density in the F2 layer at the position of the satellite.
(*a*) Bearing.
(*b*) Height.

were obtained on it up to 1145, when it was lost; there was a period from 1137 to 1141 when it was heard but was too weak for bearings to be read.

To enable the propagation phenomena to be studied further, the track of the satellite at the times in question was plotted on a map with an azimuthal equidistant projection centred on Slough, as shown in Fig. 2. In the absence of more accurate data, the orbital parameters as calculated by the staff of the Royal Aircraft Establishment[2] were used to derive the track. It was estimated that the error in bearing resulting from this procedure would not exceed 1°. The theoretical bearings measured from the map have been plotted as the solid line in Fig. 1(*a*), and it can be seen that they agree fairly well with the observed values. The departure of the points from the line between 1129 and 1136 U.T. may be significant, but it is not large and the propagation may reasonably be considered to be taking place along a great-circle path.

The next object of study was to see whether our knowledge of ionospheric propagation was adequate to explain the reception

Written contributions on papers published without being read at meetings are ited for consideration with a view to publication.
The paper is an official communication from the Radio Research Station, Depart-nt of Scientific and Industrial Research.

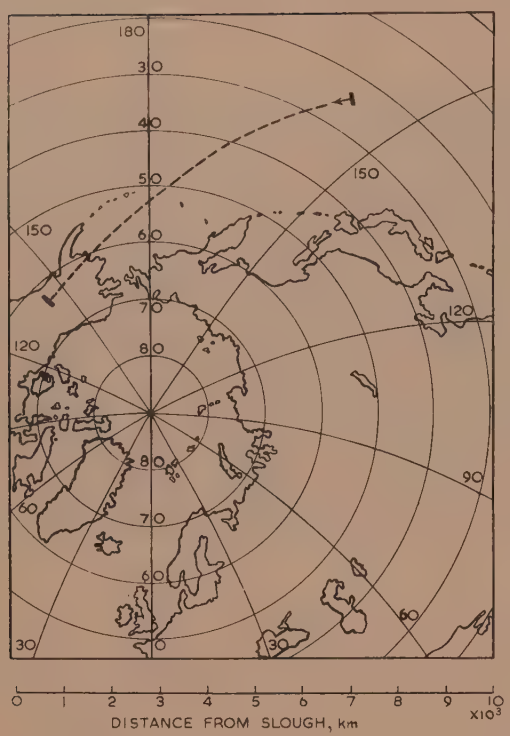


Fig. 2.—Track of the satellite 1957α on an azimuthal equidistant projection for the period 1129·5 to 1144·8 U.T. on 7th October, 1957.

of signals along these paths. First, the ionospheric predictions issued at Slough for October, 1957, were found to show that the ionization densities between Slough and the satellite were too low to permit propagation at 20 Mc/s by reflection from the F2 layer at any time in the reception period. However, ionospheric measurements at this time showed that the maximum usable frequencies (m.u.f.) in the north polar region were 10–35% higher than the predicted values. A 35% increase would be sufficient to give propagation from all points on the track—and, indeed, from points outside the portion shown. The increase was not as great as 35% over the whole region, and this may account for the period of weaker signals which occurred in the middle of the track, but ionospheric recording stations are not sufficiently closely spaced to determine whether this is the explanation.

The height of the satellite must now be considered, and a calculated curve is therefore shown in Fig. 1(*b*). It can be seen that at the beginning of the period of reception the satellite is descending through the region of maximum ionization of the F2 layer and at the end of the period it is rising through it. It seems probable that this is the factor which determines the duration of the period of reception, since the radio waves cannot, of course, penetrate obliquely through this region of high electron density.

It appears that the propagation on this occasion can be explained in terms of the usual ionospheric mechanisms. Attention is drawn to the fact that the period of reception coincides with that for which the satellite was below the F2-layer maximum; this same phenomenon can take place in closer transits,[3] at least for satellite 1957α.

(4) EFFECT OF IONOSPHERIC REFRACTION ON RECEPTION AND OBSERVED DOPPLER SHIFT

Examination of the variation of Doppler frequency with time for transits of satellites 1957α and β reveals striking discontinuities which suggest that changes were occurring in the modes of propagation of the signals. One characteristic form of Doppler curve is illustrated in Fig. 3, which is for the NW–SE transit of 1957β at 1500 U.T. on the 4th November. A weak signal was received first when the satellite was about 7000 km

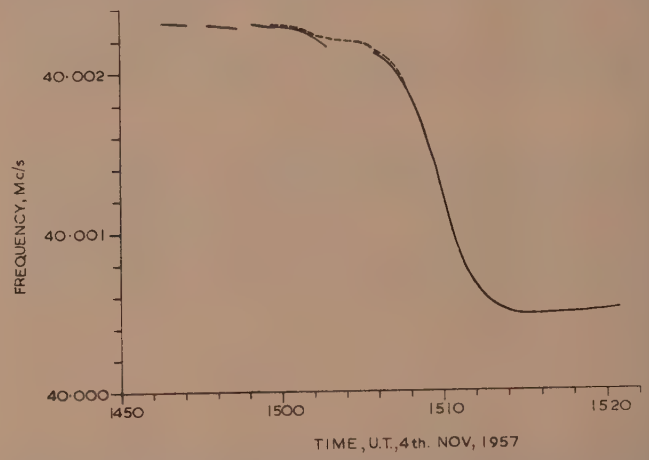


Fig. 3.—Comparison of calculated and observed Doppler curves for a north-west to south-east transit of satellite 1957 β observed at Slough on 4th November, 1957.

——— Observed.
– – – – Calculated assuming parabolic layer.

NW of Slough, and at 1502 U.T. the frequency was observed to fall as if the satellite was about to pass overhead. The signal then faded out and reappeared at 1506 U.T., when the frequency traced out a normal S-shaped curve. Since this behaviour was observed several times at Slough and elsewhere, it was decided to find out whether such a curve could be calculated from the known configuration of the orbit and measured ionospheric height and critical frequency, assuming a parabolic ionic density distribution.

The track of the satellite on this occasion passed the apex (65° N) over northern Canada at longitude 67° W, the nearest point of approach to Slough being 740 km to the south-west. Ionosphere soundings were available for the relevant time from Slough, Tortosa (Spain), Reykjavik (Iceland), Narsarssuak (Greenland) and Baker Lake and Churchill (Canada). From these it appeared that the ionization of the F2 layer was nearly constant along the satellite track, with f_0 (F2) approximately 14 Mc/s. The height of maximum ionization, h_m, was 400 km at Slough and 380 km at Tortosa, falling to 325 km at Narsarssuak and rising again to 380 km at Churchill and Baker Lake. In the calculation this small horizontal gradient was neglected and a concentric, parabolic layer with $h_m = 385$ km, $f_0 = 14{\cdot}3$ Mc/s and semi-thickness $y_m = 100$ km assumed.

Expressions were derived from parabolic layer theory (see Appendix) for the path of a ray travelling through such a layer and arriving at an angle of incidence ϕ_0 on the ground at the observer. The results are presented in Fig. 4, which shows a number of calculated ray trajectories plotted to Cartesian co-ordinates so as to permit expansion of the height scale. In the same illustration the track of the satellite has been replotted to show height against distance from Slough. This is necessary, since the satellite passed 740 km to the south-west of Slough, and the rays did not therefore propagate in the orbital plane of the satellite, or in the same plane as each other. This method of presentation permits the superposition of all the rays on one graph.

Examination of the Figure shows that, before 1458·6 U.T., '1½ hop' propagation was possible between the satellite and

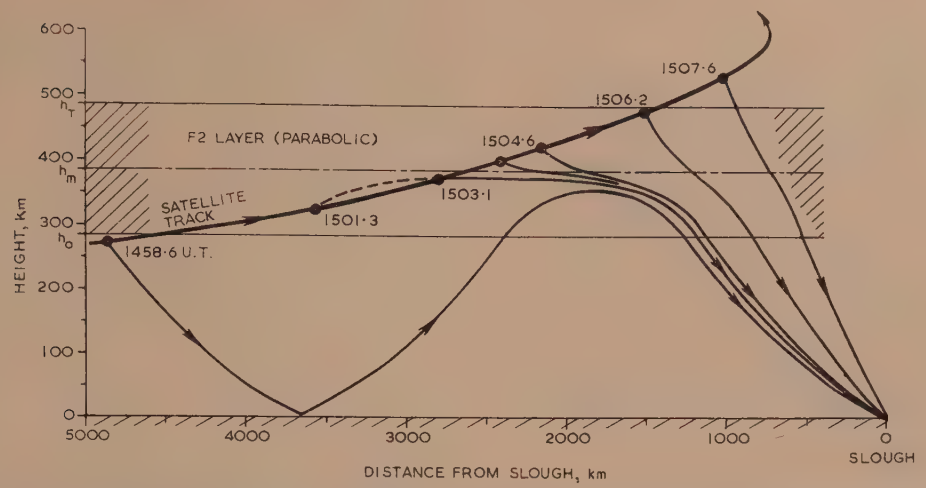


Fig. 4.—Transit of satellite 1957 β observed at Slough at 1500 h U.T. on 4th November, 1957, showing the calculated trajectories of radio rays reaching Slough from the satellite at the indicated times.

Note that the curved-earth configuration is plotted in Cartesian co-ordinates with an expanded height scale.

Slough. The ray shown for 1458·6 represents the limiting ray for this mode. After this time propagation took place by high-angle or Pedersen rays of the type shown for 1501·3. When the satellite had risen to a level in the layer slightly below the height of maximum ionization, rays of the 'escape' type became effective, as shown for 1504·0. Henceforth the rays became progressively less deviated until the satellite reached its nearest point of approach to Slough. The period for which the satellite receded from Slough is not shown in the Figure; the satellite was then above the layer and escape rays were thus effective throughout.

It will be noticed that the observed fade-out from 1503–1506 U.T. corresponds to the time for which the satellite was rising through the maximum and upper half of the layer. The theory does not account for such a fade-out of signals, although considerable defocusing and thus reduction of signal strength would be expected near the time of change-over from reflected rays to escape rays.

The theoretical Doppler shift has been calculated by means of a formula due to Weekes,[4] and is given in the Appendix. This gives the Doppler shift as a function of the known velocity and position of the satellite, and of the refractive index μ_s, at the satellite and the angle of incidence, ϕ_0, at the observer, these last two being quantities which can be derived from the parabolic-layer equations.

The resulting calculated curve for the period 1459–1508 U.T. is shown in Fig. 3. It will be seen that the frequency falls at about 1502 U.T. below the value anticipated for free-space propagation, and remains nearly stationary until 1505 U.T. This behaviour can be understood if it is realized that the effect of ionospheric refraction is to make the angle of incidence at the observer change more slowly as the satellite passes through the layer maximum, resulting in a reduced Doppler shift. As shown in the Appendix, another factor is the increased wavelength in the ionosphere; this results in a decreased rate of addition of wavelengths—another way of regarding Doppler shift. The theoretical curve is similar to the observed curve, apart from the fade-out effect. This may be attributable to the horizontal gradient in the layer, which has been neglected in this analysis.

The interesting aspect of this analysis from the aspect of radio-propagation studies is the reception of Pedersen rays having very long hops. Other evidence[5, 6] for such trajectories has recently been given, and these studies may lead to more consideration of the effect in the investigation of long-distance communication.

One other feature of satellite Doppler curves which has attracted attention is the slight rise in frequency observed as the satellite recedes. An example of this effect can be seen in Fig. 3. It appears that the decrease in the satellite's velocity (by 50 m/sec/min) as its height above the earth increases during this period results in a decreased Doppler shift sufficient to account for the observed effect.

(5) MEASUREMENTS TO HELP IN ORBITAL DETERMINATION FOR THE SATELLITE 1958 δ

An example of the measurement of bearing, angle of elevation and Doppler frequency made on the satellite 1958 δ is shown in Fig. 5. Parameters defining the size, shape and orientation of the orbit can be readily deduced from measurements on three nearby transits, two north-going at height h_1, say, and one south-going at height h_2. The two transits at h_1 are required to give a value for the period; successive orbits can be used if a rough value only is required, but a more accurate result is obtained by using two transits spaced by about 24 hours. The inclination of the orbit can be deduced from bearing measurements on any of the nearby transits, and the other parameters obtained from a calculation of satellite position at a point in its track during its transits at the two different heights; these positions can be derived,[7] for instance, from combined measurements such as those in Fig. 5.

To give some idea of the accuracy which can be achieved when using these observations for orbital calculations, some figures derived from the results in Fig. 5 alone are given in Table 1. They are compared with values deduced from some orbital details issued by the Royal Aircraft Establishment at an early stage in the life of the satellite.

Table 1

PARAMETERS DEDUCED FROM MEASUREMENTS SHOWN IN FIG. 5

	From Slough observations	*From R. A. E. orbit*
Inclination, deg. ..	66	64·9
Height, km ..	180	222
Velocity, km/sec ..	7·4	8·2

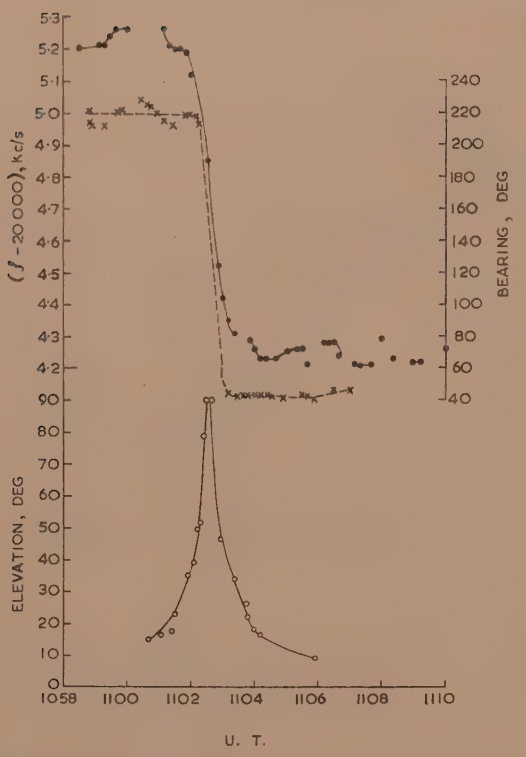


Fig. 5.—Observations of frequency, f, bearing and angle of elevation on the 20 Mc/s transmissions from the satellite 1958 δ on 18th May, 1958.

● Frequency.
× Bearing.
○ Angle of elevation.

It can be seen that the accuracy is sufficient to provide an approximate orbit, which may be very useful in the first days of the life of a satellite.

The transit in Fig. 5 is of particular interest, since it is an overhead one. The oscillations in apparent bearing during the approach are instrumental; the fluctuations are unusually large and are possibly associated with radiation of only a small vertically polarized component towards the direction-finder at the time. The reasons for the irregularities and discontinuities in the shape of the Doppler curve have not yet been investigated. They appear to be more extensive than those for satellite transmissions in the winter, and may be associated with the more complex layer structure in the summer daytime.

(6) CONCLUSIONS

Some examples have been given of the ionospheric information which has been obtained with the aid of the transmissions at 20 and 40 Mc/s from the Russian earth satellites. The high speed of the satellite over the surface of the earth enables ionospheric conditions to be sampled over widely separated parts of the earth before any large-scale changes in the ionosphere can occur. For instance, the reception of 20 Mc/s signals for 15 min when the satellite was over the Pacific Ocean has confirmed the presence of ionization densities in the north polar region higher than those predicted, as was also shown by ionospheric soundings. Again, the study of the modes of propagation on a nearby transit at 40 Mc/s has shown that the existence of a Pedersen ray must be taken into account more frequently than has perhaps been the case in the past.

The results of the Doppler analysis have helped to explain two prominent features of the observed frequency/time curves which have previously attracted comment. The first concerns the discontinuity between the early portion of the Doppler curve observed at times during the approach of the Russian satellites and the main S-shaped curve accompanying overhead transit. This change in frequency is too large to be explained by a change of mode of propagation, for instance from '½ hop' to '1½ hop', and appears to correspond with the passage of the satellite through the maximum of the F2 layer. The second result concerns the rise in frequency at the end of the observed Doppler curve; this can be accounted for by the retardation of the satellite as its height increases after perigee.

For researches of the kind described above to be possible, an accurate knowledge of satellite position as a function of time is essential. The observations made at Slough can be made to yield an approximate orbit, as has been pointed out here; this may well be useful in the early days of the life of a satellite, but no radio observations at 20 or 40 Mc/s appear likely to give a really good orbit, since, for instance, the ionospheric effects are too great. The accuracy required is ±10 km in position and ±1 sec in time for radio-propagation investigations of the type discussed in the paper, although in other radio work an accuracy at least ten times greater than this would be necessary. In either case optical observations would have to be used, and are indeed of the greatest importance in getting the most from the radio measurements.

The fact that the Russian satellites have in the first month of their life moved up and down through the F2 layer during each orbit has been of considerable assistance in research into the nature of the ionosphere. However, for the study of propagation, transmissions from a satellite which kept at a constant height below the F2-layer maximum would be very useful. It should be noted that most, if not all, satellites will be in such an orbit towards the end of their life.

(7) ACKNOWLEDGMENTS

The authors wish to acknowledge their indebtedness to their colleagues at the Radio Research Station who participated in the observational work and analysis; in particular, Mr. M. D. Vickers carried out many of the calculations for Section 4.

The work described above was carried out as part of the programme of the Radio Research Board. The paper is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(8) REFERENCES

(1) Wilkins, A. F., and Minnis, C. M.: 'Arrival Angle of H.F. Waves', *Wireless Engineer*, 1956, **33**, p. 47.

(2) Staff of the Royal Aircraft Establishment: 'Observations on the Orbit of the First Russian Satellite', *Nature*, 1957, **180**, p. 937.

(3) Bain, W. C., and Meadows, R. W.: 'Observations of Bearing and Angle of Elevation of Satellite 1', *Proceedings I.E.E.*, 1958, **105** B, p. 91.

(4) Weekes, K.: 'On the Interpretation of the Doppler Effect from Senders in an Artificial Satellite', *Journal of Atmospheric and Terrestrial Physics*, 1958, **12**, p. 335.

(5) Warren, E., and Hagg, E. L.: 'Single-Hop Propagation of Radio Waves to a Distance of 5 300 km', *Nature*, 1958, **181**, p. 34.

(6) Kift, F.: 'Single-Hop Propagation of Radio Waves to a Distance of 5 300 km', *ibid.*, p. 1459.

(7) Kitchen, F. A., Joy, W. R. R., and Billam, E. R.: 'Some Simultaneous Radio Direction-Finding and Doppler Frequency Measurements on the Russian Artificial Earth Satellites', Fifth Meeting of C.S.A.G.I., Moscow, 1958.

(8) APPLETON, E. V., and BEYNON, W. J. G.: 'The Application of Ionospheric Data to Radio-Communication Problems', *Proceedings of the Physical Society*, 1940, **52**, p. 518.

(9) APPENDIX

The Doppler shift in frequency of the signal received from an earth satellite is given by

$$f_d = \frac{d}{dt}\left\{\begin{array}{c}\text{number of wavelengths in the ray}\\ \text{path from observer to satellite}\end{array}\right\}$$

$$= \frac{d}{dt}\int_{path}\frac{ds}{\lambda/\mu}$$

where λ is the free space-wavelength and μ the refractive index of the medium. If the ionosphere is spherically stratified and concentric with the earth, we may use Bouger's rule; this relates the angle, ϕ, between the ray and the radius, r, from the centre of the earth to the refractive index:

$$\mu r \sin\phi = \text{constant} = r_0 \sin\phi_0 \quad . \quad . \quad . \quad . \quad (1)$$

r_0, ϕ_0 are the values at the surface of the earth where the observer is situated ($\mu_0 = 1$). For this condition Weekes[4] has shown

$$f_d = \frac{1}{\lambda}\frac{d}{dt}\int \mu ds = \frac{1}{\lambda}\left[r_0\dot{\theta}\sin\phi_0 + \dot{r}_s\sqrt{\left(\mu_s^2 - \frac{r_0^2}{r_s^2}\sin^2\phi_0\right)}\right] . \quad (2)$$

where θ is the angle subtended at the centre of the earth between the observer and the satellite and r_s is the radial distance of the satellite from the centre of the earth. This enables the Doppler shift to be calculated if the velocity and position of the satellite are given and if the variation of the refractive index with r is known.

To derive the theoretical curve shown in Fig. 3, the relationship between μ_s and ϕ_0 and the known quantities r and θ was found by assuming a parabolic layer having

$$\mu^2 = 1 - \frac{2}{x^2}\frac{y}{y_m} + \frac{1}{x^2}\left(\frac{y}{y_m}\right)^2$$

where y_m = Semi-thickness of layer.
y = Depth of radial penetration of ray into layer.
$x = \dfrac{\text{Transmitted frequency}}{\text{Critical frequency}}$

With this value for μ, the ray path $r = f(\theta)$ may be found from[8]

$$\theta = \int \frac{r_0 \sin\phi_0 dr}{r\sqrt{(\mu^2 r^2 - r_0^2\sin^2\phi_0)}}$$

Although eqn. (2) is in a convenient form for calculation, a clearer physical understanding of its meaning may be obtained by a rearrangement using the identity (1).

Thus $$f_d = \frac{\mu_s}{\lambda}(r_s\dot{\theta}\sin\phi_s + \dot{r}_s\cos\phi_s)$$

The term in brackets is the component (v_{ray}, say) of the velocity of the satellite along the direction of the radio ray leaving the satellite for the observer.

Thus $$f_d = v_{ray}/\lambda_s \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where $\lambda_s = \lambda/\mu_s$ is the wavelength in the ionosphere at the satellite.

If we consider the related problem of transmission of waves from the location of the observer and reception at the satellite, eqn. (3) has a straightforward interpretation. It represents the component of the satellite velocity normal to the wavefront at the satellite divided by the wavelength in the medium, and is therefore the number of extra waves traversed per unit time, which is clearly the Doppler frequency. If the effect of the earth's magnetic field can be neglected, consideration of the reciprocity theorem shows that if the expression is true for this problem it is also true for transmission from a moving satellite. Eqn. (3) is therefore general and not restricted to a spherically stratified layer, as is eqn. (2).

621.382

The Institution of Electrical Engineers
Paper No. 2914
May 1959
©

TRANSISTORS AND ASSOCIATED SEMICONDUCTOR DEVICES

A Review of Progress

By R. G. HIBBERD, B.Sc., Associate Member.

SUMMARY

After a brief introduction, some statistics on the growth of the semiconductor industry since 1948 are quoted. Mention of the main operations involved in the processing of germanium and silicon is followed by a general discussion of semiconductor-device technology, including *p–n* junction formation and other aspects of device manufacture. Reference is made to publications on various aspects of semiconductor device theory, which is followed by a detailed account of the types and characteristics of the various semiconductor devices now commercially available, including diodes and rectifiers, junction transistors and photo-electric devices. Finally, brief mention is made of the main fields of application of diodes and transistors.

(1) INTRODUCTION

The group of devices to be discussed in the Review is sometimes known in this country as 'crystal valves'. In 1946 a paper with that title was read before The Institution by Bleaney, Ryde and Kinman.[1] In 1951, two further papers, 'Crystal Diodes'[2] by Douglas and James and 'Crystal Triodes'[3] by Scott, were presented. More recently a further paper[4] by Scott was also called 'Crystal Valves'. There is much to be said in favour of such a simple collective title. Another possibility, when one considers the use of these devices alongside, or instead of, thermionic valves, is the term 'semiconductor valves'. However, the word 'transistor' has become so well established internationally that it is now customary to include it in any group title referring to the family of semiconductor devices, which results in rather lengthy titles such as that of this Review. Whether a simple collective title covering semiconductor devices in general will ever be adopted in the future remains to be seen. There are obvious international difficulties.

(1.1) General

Until the development of the transistor and subsequent *p–n* junction technology, semiconductor devices were confined to point-contact diodes, metal rectifiers and photocells, which were all constructed around a metal-to-semiconductor contact. Metal rectifiers, which have been used extensively for power rectification since 1925, are considered as a separate class, and are not included in the Review. The point-contact diode, although ostensibly of metal-to-semiconductor construction is included, partly because it can be considered as the starting point of transistor development and partly because some modern theories of its operation postulate the presence of a *p–n* junction in the semiconductor after the forming process.

There is little doubt that the transistor has been the most significant addition to the field of electronics since the thermionic valve. Even in its original form—the point-contact transistor—it indicated the possibility of reducing the size, weight and power consumption of many types of electronic equipment. This early type, however, suffered from inherent limitations in power-handling capacity and operating frequency. Furthermore, it had a high noise figure and was liable to instability under certain load conditions. As a result, it had a relatively short life, and can now be considered as obsolete. Therefore it is not proposed to discuss the point-contact transistor further in the Review. The junction transistor has overcome most of the limitations of the point-contact type and is now becoming available for use over a wide range of powers and frequencies, which is being continually extended.

The intensive research and development which followed Shockley's publication on the junction transistor soon found an important outlet in the junction rectifier. In view of its simpler nature it was possible to complete developments and get designs into production more quickly than with transistors. A wide variety of germanium and silicon rectifiers soon became commercially available at reasonably economic prices. They show every sign of completely replacing copper-oxide, selenium and thermionic rectifiers in the electronic field, and are also proving serious competitors to mercury-arc and other rectifiers in the high-power field. We can thus see the significance of the transistor in relation to the semiconductor rectifier. The discovery of the transistor, together with the subsequent intensive development work, made possible the rapid advances in junction-rectifier technology.

Owing to the development effort being concentrated, for commercial reasons, on transistors and rectifiers, utilization of the optical and thermal properties of semiconductors has been, in comparison, somewhat slow. Some developments have, however, taken place, and a few photo-diodes and photo-transistors are available. One of the most significant developments in this field has been that of the silicon solar cell, which promises to have a great future if the economics of production can be satisfactorily solved.

In the Review, an attempt has been made to cover the whole subject of semiconductor devices, from starting material to applications, in a very general way. When we consider that, up to the present, some 3 000 articles and about 20 books have been written on the subject, it is obvious that it is not possible to treat every aspect in the detail it may appear to merit, or even, in every case, to include mention of all the facets of a particular development.

It is of interest to note that a journal devoted entirely to references and abstracts of technical papers on semiconductor topics is now published.[5] In its main classification index it lists 600 subject headings.

The references in the Review are not intended to be exhaustive. They list articles which are particularly appropriate to the context, or are generally considered by those working in the field to be of a 'classic' nature.

(1.2) The Growth of the Semiconductor Industry

The growth of the transistor and its associated devices has been quite remarkable. Table 1 illustrates how developments have taken place in quick succession since the invention of the transistor in 1948, and indicates the rapid growth of research and development effort. It is interesting that, of all these devices, the only one now obsolete is the original point-contact transistor.

The Table shows that transistors, covering on the one hand

Mr. Hibberd is with the British Thomson-Houston Co., Ltd.

Table 1

LANDMARKS IN SEMICONDUCTOR-DEVICE DEVELOPMENT

Year	Material	Device	Special feature
1940	Silicon	Radar crystal diode	3 000 Mc/s
1945	Germanium	Point-contact diode	100 volts
1948	Germanium	Invention of the transistor (point contact)	Amplification
1950	Germanium	Grown-junction transistor	100mW
1950	Germanium	Photo-diode and photo-transistor	High sensitivity
1951	Germanium	Alloy-junction rectifier	0·5 amp
1952	Germanium	Alloy-junction transistor	200 mW
1953	Germanium	Grown-junction tetrode transistor	50 Mc/s
1953	Germanium	Alloy power transistor	20 watts
1953	Silicon	Alloy-junction diode	0·1 amp
1953	Silicon	Zener diode	Low-voltage reference
1953	Germanium	Double-base diode (unijunction transistor)	Trigger action
1953	Germanium	High-power rectifier	50 amp
1953	Germanium	Surface barrier transistor	50 Mc/s
1954	Germanium	*p–n–i–p* transistor	400 Mc/s
1954	Germanium	Alloy power transistor	100 watts
1954	Germanium	Rate grown transistor	10 Mc/s
1954	Silicon	Grown-junction transistor	150° C
1954	Silicon	Solar battery	High efficiency
1955	Germanium	Melt-back transistor	10 Mc/s
1955	Germanium	Field-effect transistor	Unipolar
1956	Silicon	V.H.F. diffused-junction transistor	100 Mc/s
1956	Germanium	V.H.F. diffused-junction transistor	500Mc/s
1956	Silicon	Low-power junction rectifier	600 volts
1956	Germanium	Alloy drift transistor	30 Mc/s
1957	Germanium	Post-alloy diffused transistor	100 Mc/s
1957	Germanium	Spacistor	3 000 Mc/s
1957	Silicon	High-power junction rectifier	150 amp
1957	Silicon	Diffused-junction power transistor	80 watts
1958	Germanium	Technetron transistor	500Mc/s
1958	Silicon	Controlled rectifier	10 amp pulse triggered
1958	Silicon	H.V. low-power rectifier	1·5 kV
1958	Germanium	Very-high-current rectifier	500 amp

powers up to 100 watts, or on the other frequencies of up to 3 000 Mc/s, have been achieved, as have rectifiers with currents of up to 500 amp or reverse voltages of up to 1·5 kV. Several special devices have also been initiated, such as the solar battery, and the controlled rectifier.

At present, apart from work in universities and government laboratories all over the world, there are well over 80 commercial companies engaged on the development and production of semiconductor devices—about 50 in the United States, 18 in the United Kingdom, 10 on the continent of Europe, 8 in Japan, and some effort building up in Canada and Australia. Although, as in other spheres, we have little knowledge of activities in the Soviet Union, there can be little doubt that considerable effort is also being expended there.

Until 1953, production was only very small, and mainly of an experimental nature. The subsequent build-up has been progressive. In 1958, about 50 million transistors and 100 million diodes (including rectifiers) were made in the United States, and about 4 million transistors and 6 million diodes in this country. Several estimates have recently appeared of the possible trend of future production of these devices. Typical of these are that, in 1965, the output of transistors in the United States will be about 250 million and of diodes 350 million, while, in this country, the output of each may be about 50 million. Fig. 1 compares these figures with those for thermionic receiving

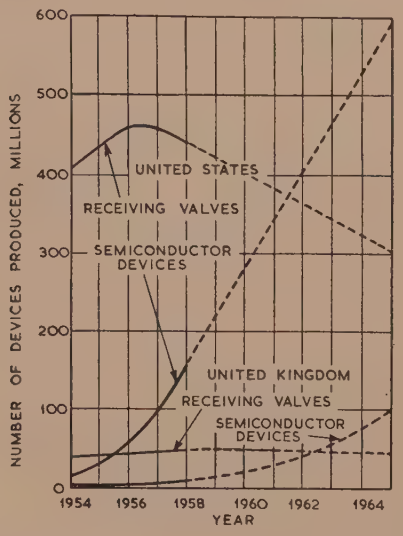


Fig. 1.—Comparison of semiconductor-device and thermionic-valve production.

valves. It indicates that, in both countries, the total output of semiconductor devices will overtake that of receiving valves by about 1962.

A further indication of the rapid growth of the industry can be seen from the fact that, in 1958, the American journal *Electronic Industries* listed 600 transistor types and 1 200 rectifier types in the United States. Although no similar lists have been issued in this country, the British manufacturers' literature indicates about 100 transistor types and 160 diode and rectifier types.

These large numbers of types have appeared because, owing to the speed of developments, it has not yet been possible to introduce agreed standardization. Each manufacturer has his own version of a particular device, and what, with standardization, could be one type, at present represents perhaps 20 slightly different types. In other words, the 600 transistor types listed in the United States relate to about 30 basic types.

In this country, progress of semiconductor work, although not quite comparable with that in the United States, is in itself quite impressive. The silicon radar crystal was pioneered in this country, and notable work was carried out on the processing of germanium from flue dust.[6] The germanium power rectifier was quickly developed, and a 300 kW unit using germanium cells was installed in 1953.[27] With transistors, mention should be made of the development of the post-alloy diffusion technique[16] and the cold-weld method of encapsulation. A recent British development of far-reaching importance is the preparation of very-high-purity silicon by a silane process. At least six or seven new factories with air-conditioning and other special features have been erected, and production is building up. So far, the production of transistors has, in general, been limited to the *p–n–p* alloy type, but other types are gradually appearing to meet the growing demand as more application work proceeds.

(2) SEMICONDUCTOR MATERIALS

Of the possible semiconductor materials, germanium and silicon are the only two at present being used for commercial devices. To be suitable for junction transistors and rectifiers the

material must be in the form of a single crystal with a certain resistivity. This required resistivity is obtained by adjusting the amount of a particular impurity. The amount needed is extremely small—for transistors approximately one part in 10^8, and for high-voltage rectifiers only one part in 10^9. To obtain such low impurity concentrations, the germanium or silicon is first refined to a higher purity—about one part in 10^{10} if possible—and the required amount of impurity is introduced during the process of obtaining a single crystal.

One of the most important developments in the processing of both germanium and silicon has been the purification by zone refining,[7] in which a molten zone is passed along a solid bar. Most impurities are more soluble in the molten than in the solid metal, and thus a succession of molten zones passing slowly along the bar conveys these impurities to one end, leaving the rest of the bar very pure. This process is extremely successful with germanium using graphite or quartz 'boats' and significant impurities can readily be reduced to the required order of one part in 10^{10}. Germanium so refined is then converted into single-crystal form by a process called the 'pulling'[9] technique. A quantity of the refined germanium, together with the requisite amount of impurity, is raised to a temperature a few degrees above its melting point in a graphite or quartz crucible. A small piece of single-crystal germanium known as a 'seed' is lowered into contact with the surface of the melt and then slowly raised at the rate of a few inches per hour. The metal solidifies on to the seed with a continuous crystal structure as it is withdrawn.

However, molten silicon is so reactive that materials, such as quartz, normally used for crucibles are slightly soluble in it and thus any impurities in the quartz are introduced into the melt. This imposes a limit to the purification that can be achieved in zone-refining silicon using quartz boats, and also to the purity of a single crystal 'pulled' from a quartz crucible. To overcome both limitations, a variation of the zone-refining technique has been developed for use with silicon. A rod of polycrystalline silicon is supported vertically, held at the top and bottom, and a molten zone, small enough to be held in position by surface tension, is formed at the bottom and then slowly passed up to the top, taking the impurities with it. If a section of single crystal is positioned at the bottom, and the molten zone passed from it to the polycrystalline rod, the latter will be converted to single-crystal form simultaneously with the refining. This process is called the 'floating-zone' process[8] and promises to be of great importance in silicon technology. At present, the majority of silicon single-crystal material is produced by the pulling method, but it would appear that, in the near future, a considerable amount will be produced by this floating-zone method.

The present limit of purity of silicon appears to be the residual content of boron—an acceptor impurity. Since boron has a segregation constant of approximately unity, it is extremely difficult to remove by zone refining. As a result, considerable effort has been made to remove it by chemical means, and in this connection, a noteworthy advance has recently been the preparation of silicon by a silane process.[10] By this means, it is possible to reduce the boron content to less than two parts in 10^{10}.

Having obtained a single crystal, evaluation measurements are necessary. Details of these are well known, and it is sufficient to state that resistivity, minority-carrier lifetime, dislocation density and crystal orientation are all important from the device viewpoint and must be held under control. Table 2 gives typical material specifications for various devices.

To prepare the material in a form suitable for device fabrication it must be cut up. The crystal is first sliced into wafers about 0·015 in thick, and then each slice is cut into rectangular pellets varying in size from about $\frac{1}{16}$ in square upwards, depending upon the device to be made. Three main methods of cutting are used—diamond-wheel cutting, wire cutting and ultrasonic cutting. Each manufacturer has his own preference in this matter, but whatever method is used, he must be very conscious of the need to reduce the cutting waste as much as possible. The value of single-crystal silicon is about 15s./g, which, in engineering units approaches £1 million per ton.

Table 2

TYPICAL MATERIAL SPECIFICATIONS

Device	Type	Starting material	
		Resistivity	Lifetime
		ohm-cm	microsec
Silicon Zener diode	*n*	0·03	—
Germanium h.f. alloy transistor ..	*n*	0·5	25
Germanium a.f. alloy transistor ..	*n*	2·0	50
Germanium power alloy transistor	*n*	6·0	100
Germanium power rectifier ..	*n*	30	1 000
Silicon power rectifier	*n*	150	100
Silicon power rectifier	*p*	1 000	1 000

(3) SEMICONDUCTOR-DEVICE TECHNOLOGY

Both junction rectifiers and junction transistors consist of a combination of one or more *p–n* junctions together with ohmic contacts to the various semiconductor regions.

Although obvious emphasis must be placed on the *p–n* junction from the operational characteristics of any junction device, the ohmic contacts are also very important, and they are often quite difficult to execute in practice. The various methods developed for making *p–n* junctions and completing device structures through to the final encapsulated device have resulted in quite a new technology that may best be reviewed by considering the various operations in turn.

(3.1) *p–n* Junction Formation

Many alternative methods of producing *p–n* junctions have been evolved. In a recent comprehensive paper,[11] Dacey and Thurmond describe some 12 methods, which fall into three basic groups—grown, alloyed, and diffused junctions.

(3.1.1) Grown Junctions.

The first method developed for producing a *p–n* junction was to add the correct amount of an opposite type of impurity to the melt at a given time during the growth of a single crystal. This converts the growing crystal from *p* to *n*, or vice versa, at that point.

For transistor structures a small quantity of the original impurity is then added after a few seconds to produce a second junction. Thus the *p–n* structure, either single or multiple, is grown directly in the single crystal, which is subsequently cut into small bars so that each includes the *p–n* structure. Although considerable numbers of transistors have been produced by this 'doped-grown' process, the method is tending to lose favour because of the poor economics of production and also because of the limitations introduced by the inherent high series resistance of the collector and emitter regions.

Several variations of grown junctions have been developed for transistor structures, depending upon segregation effects on solidification.[12] Here a homogeneous crystal is grown with both donor and acceptor impurities, and is cut into small bars. If one end of a bar is melted and allowed to solidify, the relative initial concentrations of the donors and acceptors in the recrystallized region will depend upon the segregation constants of the impurities, and by suitable choice, a *p–n* junction or a thin

p-layer may be formed. Such processes have been called 'melt-back', 're-melt' and 'melt-quench', and have economic advantages over the 'doped-grown' method in that virtually the whole of the single crystal is usable instead of only a small fraction. Another variation, in which this melting back takes place during the growth of the crystal is called 'rate growing'. This again has the advantage that many junctions can be produced in one crystal as opposed to only one or two for the doped-grown method.

(3.1.2) **Alloy Junctions.**

The alloy method of forming a *p–n* junction, originally introduced by Hall and Dunlap,[13] has proved to be a most important process. In it, an opposite type of impurity is alloyed into one surface of a thin wafer of the semiconductor material. Perhaps the most common combination is indium alloyed into *n*-type germanium. During the alloying operation, the temperature of the combination is raised to about 600° C for a short time, and on subsequent solidification the germanium recrystallizes heavily doped with indium to form a *p*-type layer adjacent to the unmelted *n*-type material. It has proved possible to develop this method for large-scale production, the alloying cycle being carried out either as a continuous process using a moving-belt tunnel furnace, or in a static muffle furnace with multiple jigs to alloy up to 200–300 units at a time.

(3.1.3) **Diffused Junctions.**

The most recently developed, and perhaps the most important, method of forming *p–n* junctions is by diffusion of opposite-type impurities into the solid semiconductor.[14] Suitable wafers of the semiconductor are heated to a temperature approaching the melting point in a gaseous atmosphere containing the impurity. Typical temperatures are 850° C for germanium and 1 250° C for silicon. The impurity atoms diffuse into the solid crystal, and, under the correct conditions, they will convert the surface layer to the opposite type of conductivity and form a *p–n* junction just inside the crystal. The process is very slow. For instance, to diffuse, say, phosphorus into silicon to a depth of 0·001 in, diffusion at 1 200° C for about 15 hours is necessary. This means, however, that diffusions of such depths can be controlled very accurately, and also that very small penetrations—down to a few millionths of an inch—are possible with short diffusion times. Since a large number of wafers can be diffused in one operation, it would appear that the slowness is no obstacle to large-scale production.

(3.1.4) **Mixed Junction Structures.**

Since a transistor requires at least two junctions, and the required characteristics of these junctions are generally quite different, combinations of the different methods of junction formation may be useful for some transistor structures. Typical examples are grown-diffused transistors,[15] post-alloy diffused transistors,[16] *p–n–i–p* transistors[17] (alloyed diffused) and diffused base-alloyed emitter transistors.[18] More detailed mention of these will be made later.

One problem still outstanding in all methods of junction formation is that some change to the original crystal takes place during the heat cycle involved. This is much more serious in diffusion, owing to the higher temperatures used. However, considerable effort is being expended on detailed investigation of these effects, and in the meantime due account is taken of any change in the device design.

(3.2) Ohmic Contacts

An ohmic contact must serve as a means of getting current into or out of the semiconductor without playing any part in the active processes of emission and collection. This condition can usually be met by including in the solder, with which the contact is made, a sufficient quantity of the impurity responsible for the conductivity of the semiconductor. Thus, for contact to an *n*-type semiconductor, it is usual to include a small amount—about 1%—of a group 5 element such as antimony or arsenic, and to a *p*-type semiconductor, a group 3 element such as indium or gallium. Contacts of this type are known as *nn*+ and *pp*+ junctions.

An attractive and 'clean' technique recently introduced is thermo-compression bonding.[19] Satisfactory contacts can be made between gold and evaporated layers of aluminium or gold with pressures of a few thousand pounds per square inch at a temperature of a few hundred degrees centigrade. The process would appear to have considerable advantages over other methods for small-area contacts, and recently it has been claimed that areas of up to 0·1 in diameter are possible.

(3.3) Surface Preparation

Surface treatment of the semiconductor is usually essential both before and after junction formation—before to remove the surface layer damaged by cutting and lapping,[20] and after to remove contamination bridging the junction. This is normally carried out by chemical or electrolytic etching, and must be followed by thorough washing with deionized water, to remove all traces of the etchant. For this purpose a continuously-circulating water system[21] has recently been described. Ultrasonic vibration during this washing has been proposed, and may be beneficial in removing physical particles such as lint and dust.

(3.4) Encapsulation

Not the least of the problems associated with the development of semiconductor devices has been mounting and sealing the device structure into its envelope. Initially, the importance of hermetic sealing and the maintenance of surface conditions was not fully appreciated. The fact that the active processes in a semiconductor device take place inside a solid structure gave rise to the feeling that the relatively cumbersome vacuum techniques associated with the sealing of thermionic valves were not necessary, and many early devices were 'cast' into a plastic encapsulation or sealed with a plastic adhesive. Diodes and transistors so manufactured were very economic in production and gave fairly satisfactory service for reasonable periods under average ambient conditions, but operation in atmospheres subject to high humidity brought deterioration of characteristics owing to the ingress of moisture. Thus it gradually became apparent that hermetic sealing is essential, and encapsulation has tended towards glass envelopes for low-power devices and metal envelopes with glass-to-metal seals for higher-power devices. Evacuation and back filling with a dry inert gas have been used, but final sealing in an inert atmosphere is more common. Molecular sieves are used to some extent as getters for residual water vapour. A good discussion on the general subject of encapsulation is given in Reference 22.

There are obviously many possible variations of design that can fulfil the requirements of encapsulation satisfactorily, and no one design is necessarily more correct than another. Different manufacturers naturally prefer different techniques to suit their own preference and experience. Brief mention of a few of the alternative techniques used commercially will help to complete the picture.

(3.4.1) **Hermetic Sealing.**

For point-contact diodes, and some low-power transistors, all-glass construction is used with glass-to-Dumet seals. The

main problem is to reduce the temperature of the device element to avoid damage while the final glass-to-glass seal is made.

With the vast majority of devices, the final hermetic seal is made by a metal-to-metal joint. There are four main methods of sealing—soft solder, projection welding, argon-arc welding and cold-pressure welding. Soft solder is still widely used, the essential requirement being clean pre-tinned components with no flux residues. Projection welding is attractive, and being very suited to large-scale production, is also widely used, but it depends upon precision components with close tolerances. Argon-arc welding can be very successful, but it is not used to any great extent. Cold-pressure welding has the great advantage that the operation is cold, which minimizes the possibility of contamination, but it is limited to a few materials, copper and aluminium being the two most suitable.

A survey of the devices marketed by 25 leading manufacturers in the United Kingdom, the United States and Europe has indicated that 50% of the companies use soft solder, 60% use projection welding, 10% use argon-arc welding and 25% use cold-pressure welding. (The total of these percentages exceeds 100, as several organizations use more than one method.)

(3.4.2) Junction Protection.

To avoid contamination during sealing is always difficult, and necessitates continual care and cleanliness. In order to alleviate this problem, some manufacturers cover the device structure with a coating of protective material before the hermetic seal is made. The most common materials used for this purpose are silicone greases for germanium devices and silicone varnishes for silicon devices. There is much divided opinion on the advisability of this process, on the basis that the protection may only be temporary, and contamination may eventually diffuse through to the junction and cause failure. Amongst manufacturers, opinion is fairly evenly divided, at least on current production. A recent survey indicated that about 40% of manufacturers are using internal junction protection on some of their present devices.

(3.4.3) Standardization.

It is obviously preferable to establish some measure of standardization of electrical characteristics, outline dimensions and lead-wire arrangements. In the United Kingdom this matter is under detailed consideration by the British Standards Institution and the British Radio Valve Manufacturers' Association, and recommendations will undoubtedly be made in the near future. Initial steps in the United States have already been taken, with the introduction of the JETEC30 encapsulation.

(3.5) General Aspects of Manufacture

As stated earlier, the purity of germanium and silicon for semiconductor-device manufacture is extraordinarily high. For transistors, the amount of impurity introduced into the refined semiconductor material is only one part in 10^8, and into the rectifiers about one part in 10^9. It has been mentioned that only a trace of moisture or contamination on the active surface can seriously degrade the performance of a device. Thus it is essential that the whole manufacturing process, from the preparation of the component parts through to the final encapsulated device, be carried out under extremely clean and dry conditions. Most manufacturers have now set up extensive air-conditioned dust-free buildings, and, in addition, carry out many of the fabrication processes in enclosed benches of the 'dry box' type. Operators wear special clothing and gloves of non-linting material to minimize contamination from handling, and all subsidiary materials, chemicals and gases must be of a high standard of purity.

Since the production of standard types of transistors and rectifiers will eventually run into many millions per year, mechanization of both device fabrication and factory testing will gradually be introduced. Some manufacturers have already installed automatic equipment for the more straightforward processes and tests.

(4) SEMICONDUCTOR DEVICE THEORY

A convenient starting point in reviewing the highlights of semiconductor-device theory is Shockley's original paper[23] on the *p–n* junction in 1949. Since then, there has been considerable increase in the understanding of the extension of Shockley's ideal model to more practical structures, and consequently in the understanding of the various types of junction devices currently being developed and manufactured. A comprehensive review of the advances in the understanding of the *p–n* junction triode has recently been presented by Pritchard.[24] The present review will, of necessity, be much briefer.

The forward characteristics of a *p–n* junction have been studied by Hall,[25] Saby[26] and Blundell,[27] and the inverse breakdown characteristics have been shown to be due to an avalanche mechanism rather than to a field-emission effect, by McKay[28] and Miller.[29] Two of the first notable papers on the junction transistor were by Early,[30, 31] the first on the effects of space-charge-layer widening and the second on the general design theory of junction transistors. The decrease in current gain with increasing emitter current was shown by Webster[32] to be the result of high-level injection, and the high-frequency operation of the junction transistor was covered in two papers by Pritchard.[33, 34] With the diffusion process and non-uniform base regions came a theoretical study[35] by Kroemer on the drift transistor. The general subject of diffusion had, of course, been extensively treated in textbooks, such as that by Barrer,[36] and an early paper introducing the subject to semiconductors was one by Dunlap.[37] A classic paper on the switching characteristics of transistors was presented by Ebers and Moll[38] in 1954, and a comprehensive treatment of the power transistor by Fletcher[39] appeared in 1955. The avalanche transistor was first treated by Kidd *et al.*,[40] and four-region *p–n–p–n* structures have been the subject of papers by Moll *et al.*[41] and Aldrich and Holonyak.[42] Various theories on the subject of noise in junction transistors are presented and reviewed in a recent paper by Van der Ziel.[43]

The papers quoted are but a few of the many hundreds of theoretical presentations that have appeared on the subject, and with the ever-widening field of semiconductor devices, there will be many new theories put forward, both on existing devices and on new devices yet to be evolved. Perhaps the greatest need for more understanding is on the subject of surface phenomena, which appear to limit the performance of many present devices. A number of theoretical studies on the effects of various surface states have been made by Plummer,[44] Eriksen *et al.*[45] and Kingston,[46] but more understanding of the nature of the surface states leading to possible methods of controlling them will considerably enhance the future of semiconductor devices.

(5) SEMICONDUCTOR-DEVICE TYPES AND CHARACTERISTICS

In this Section it is proposed to discuss basic device types, in order to give an overall picture of their construction, possibilities and present status. It can generally be assumed that all basic types are available in the United States. Where availability in the United Kingdom is indicated, the statement refers to devices actually being made in this country and does not include devices that can be imported through agents.

Schematic representation of basic rectifier and transistor types, indicating their general structural features is shown in Fig. 2.

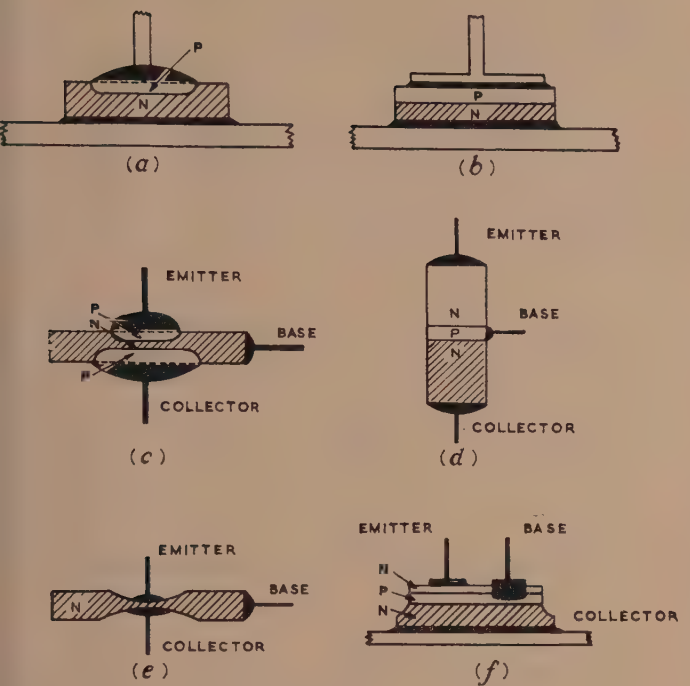


Fig. 2.—Schematic of basic rectifier and transistor construction.

(*a*) Alloy rectifier.
(*b*) Diffused rectifier.
(*c*) Alloy transistor.
(*d*) Grown transistor.
(*e*) Surface-barrier transistor.
(*f*) Diffused transistor.

(5.1) Semiconductor Diodes and Rectifiers

The use of the terms 'diode' and 'rectifier' has not been completely standardized. The author prefers to use the word 'diode' when the device is primarily intended for general circuit applications such as frequency conversion, detection, limiting, etc., and the word 'rectifier' where the device is primarily intended to rectify power supplies. Germanium and silicon diodes and rectifiers have been developed to cover a very wide range of operation, from the silicon-mixer crystal diode, handling a few microwatts at 30 Gc/s to the silicon power rectifier handling up to 20 kW per cell at mains-supply frequency.

The range of characteristics covered by the basic types is shown in Tables 3A and 3B, and Fig. 3 compares the inverse characteristics of several types.

(5.1.1) Point-Contact Diodes.

(5.1.1.1) *Silicon Crystal Diode.*

The silicon crystal diode is used almost exclusively in the microwave region[47] as a mixer or detector, mainly in radar and point-to-point communication systems. Three basic types are produced—a plug-in capsule for 3 Gc/s, a coaxial type for 10 Gc/s and a waveguide type for 30 Gc/s. All these encapsulations are designed with precision dimensions to ensure maintenance of correct matching conditions in coaxial and waveguide circuits.

(5.1.1.2) *Germanium Point-Contact Diode.*

As a general-purpose diode, covering frequencies up to several hundred megacycles per second, reverse voltages up to about 100 volts, and currents up to 10–20 mA, the germanium point-contact diode has been established for some 12 years, and many millions have been produced. This is in spite of the fact that

Table 3A

TYPICAL CHARACTERISTICS OF BASIC GERMANIUM DIODE AND RECTIFIER TYPES

Type	Maximum peak reverse voltage	Maximum mean current rating	Maximum junction temperature	Typical reverse current at maximum peak reverse voltage (25° C)	Typical forward voltage drop at maximum current	Availability in United Kingdom
	volts		deg C		volts	
Point-contact diode	100	30 mA	70	250 μA	1·2	Yes. Good
Gold-bonded diode	100	100 mA	75	10 μA	0·8	Yes. Fair
Low-power junction rectifier ..	300	0·5 amp	85	100 μA	0·5	Yes. Good
Medium-power junction rectifier ..	300	5 amp	75	500 μA	0·5	Yes. Good
High-power junction rectifier ..	300	50 amp	75	2 mA	0·5	Yes. Good
High-current junction rectifier ..	100	500 amp	75	25 mA	0·5	No

Table 3B

TYPICAL CHARACTERISTICS OF BASIC SILICON DIODE AND RECTIFIER TYPES

Type	Maximum peak reverse voltage	Maximum mean output current	Maximum junction temperature	Typical reverse current at maximum peak reverse voltage (25° C)	Typical forward voltage drop at maximum current	Availability in United Kingdom
	volts		deg C		volts	
Microwave crystal diode	3	30 mA	70	1 mA	0·8	Yes. Good
Sub-miniature diode (switching, etc.)	100	20 mA	150	0·1 μA	1·0	Yes. Fair
Miniature junction rectifier ..	400	200 mA	200	1·0 μA	1·1	Yes. Fair
Low-power junction rectifier ..	400	1·0 amp	200	1·0 μA	1·1	Yes. Good
High-voltage junction rectifier ..	1 000	1·0 amp	200	1·0 μA	1·1	Experimental samples
Zener diodes	3–12	—	200	—	—	Yes. Good
Medium-power junction rectifier ..	300	10 amp	200	20 μA	1·1	Preproduction samples
High-power junction rectifier ..	300	100 amp	200	1 mA	1·1	Preproduction samples

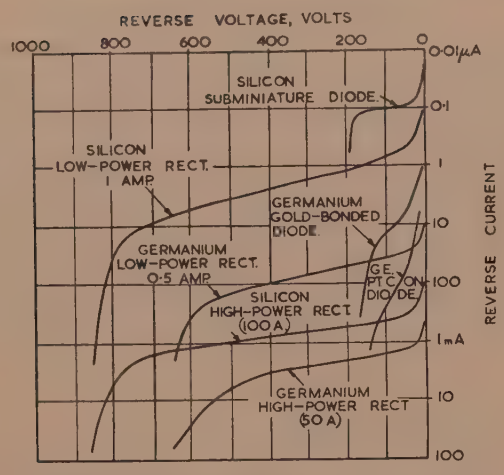


Fig. 3.—Average reverse characteristics of various germanium and silicon diodes and rectifiers.

no fully satisfactory theory of operation has yet been presented. Although there are signs that the sub-miniature silicon diode may replace it in many applications, it seems likely to continue for a considerable time if only on economic grounds.

The doping of the germanium can be controlled to give either high reverse voltage, up to 200 volts, with low-frequency cut-off, about 5 Mc/s, or a low reverse voltage, about 25 volts, with a frequency cut-off of several hundred megacycles per second. The most popular form of encapsulation at present is a small all-glass unit, approximately $\frac{1}{8}$ in diameter by $\frac{1}{4}$ in long with coaxial wire leads at the two ends.

An extensive discussion on point-contact devices is given in Reference 48.

(5.1.1.3) *Gold-Bonded Diode.*

In the gold-bonded diode, the end of a thin gold wire (0·001 5 in diameter) is bonded to the surface of an *n*-type germanium pellet by the passage of a current pulse. The gold includes a small percentage of gallium or indium, which, during the bonding, alloys into the germanium to form a rectifying contact with characteristics approaching those of a *p–n* junction (see Table 3A). It is perhaps somewhat surprising that more attention has not been paid to this diode, as its characteristics are superior to those of the point-contact diode in all respects except high-frequency performance. This is not quite so good owing to a slightly higher capacitance. Its future will depend to some extent on the progress of the sub-miniature silicon-bonded diode, but at least one manufacturer in the United Kingdom is producing it, and it may well be perpetuated as a basic type.

(5.1.2) Junction Diodes and Rectifiers.

Junction diodes and rectifiers are all based on the *p–n* junction discussed earlier in the Review. The alloy method is used exclusively with germanium, while, with silicon, favour is fairly evenly divided between alloying and diffusion.

Compared with earlier forms of metal rectifier, semiconductor junction rectifiers are remarkably efficient owing to a low forward voltage drop coupled with a high reverse voltage per cell. Reference 27 is a general paper on germanium and silicon power rectifiers.

(5.1.2.1) *Sub-miniature Silicon Junction Diode.*

The sub-miniature silicon junction diode is similar in construction to the germanium gold-bonded diode, except that a small aluminium wire is alloyed to an *n*-type silicon pellet. Characteristics show a considerable improvement over the point-contact diode—reverse voltages of 100 or 200 volts with reverse impedance over 500 megohms at 25° C and 10 megohms at 150° C. The forward voltage drop is about 1 volt at 20 mA. There seems little doubt that this diode will replace the point-contact diode for many low- and medium-frequency applications, particularly for switching purposes.

Another version of a small-area silicon junction diode, the u.h.f. 'mesa' diode recently described by Forster and Zuk,[50] uses a diffused junction with an active diameter of 0·005 in. This diode was designed for fast switching use and has switching times of the order of millimicroseconds.

(5.1.2.2) *Germanium Junction Rectifiers.*

In a typical germanium junction rectifier, indium is alloyed into *n*-type germanium to form the rectifying junction. Ohmic electrode contacts are made to the indium with a low-melting-point solder (120° C) and to the under-surface of the germanium with tin-lead alloy. The basic unit so formed is mounted into the encapsulation with a low-thermal-resistance path between the germanium and an external surface, so that external heat-radiating systems can be applied as required. The area of the *p–n* junction varies from 1 mm^2 for the 0·5 amp unit (see Table 3A) to 64 mm^2 for the 50 amp cell, the manufacturing techniques being similar for all sizes.

(5.1.2.3) *Silicon Junction Rectifiers.*

The characteristics of silicon junction rectifiers being made at present are so good that it would seem reasonable to forecast that, for all general-circuit d.c. supplies and industrial medium-power supplies, they will be accepted standard rectifiers of the future. Low-power (1 amp) rectifiers, with junction diameters of 0·025 in, can be made with working peak reverse voltages of up to 1·5 kV, the reverse current at this voltage being only a few microamperes. In the United Kingdom, the highest-voltage type commercially available at present has a reverse voltage of 600 volts, with the majority in the 300–400-volt region. A miniature silicon rectifier is now available in a glass encapsulation similar to the small point-contact diode. It uses a diffused-silicon pellet and has reverse voltages of up to 600 volts with current ratings up to 400 mA.

Details of large-area silicon rectifiers suitable for 100 amp with reverse voltages up to 1 kV have been published.[49] A few units are now available in this country with voltage ratings up to 300 volts.

The forward voltage drop of each of these silicon rectifiers is about 1 volt at its rated current. Another notable feature is the permissible junction temperature of 200° C, which allows operation with reduced current rating at ambient temperatures up to about 175° C.

As mentioned earlier, two methods of manufacture are being used—alloying and diffusion. From an overall point of view there appears little to choose between the two methods; the characteristics and selling prices of available units are very similar. The alloy method, in which aluminium is alloyed into *n*-type silicon, involves temperatures of about 650° C with relatively short time cycles, and this may appear to be more convenient than the diffusion method, which requires temperatures in the region of 1 200° C with diffusion times of several hours. On the other hand, the problem of producing uniform wetting of the silicon surface by aluminium with the alloy method is probably more difficult than that of obtaining a flat diffused junction. Extended life tests on both types of junction have given good results, and it would appear that they will both continue to be made according to the different manufacturers' preferences.

One point in favour of diffusion is the fact that the technique

can be readily applied to the whole family of diode devices, from large-area power rectifiers down to the tiny u.h.f. 'mesa' diodes. To some extent, this is also true for alloying, but the actual process details are much more dependent on the junction area.

(5.1.2.4) *Silicon Zener Diode.*

The Zener diode[51] is a special version of the silicon junction diode with a well-defined sharp reverse breakdown at low voltage. It can be operated in this breakdown region as a voltage reference or used as a limiter in much the same way as a gas discharge tube. However, it has the advantage that any voltage down to about 2 volts is possible. The slope resistance and temperature coefficient of breakdown are both dependent on the actual value of the breakdown voltage, the former being a minimum at about 6 volts and the latter passing through zero at about 5 volts. A typical low-power Zener diode with a breakdown at 6 volts will have a slope resistance of about 2 ohms at 20 mA and a temperature coefficient of breakdown voltage of +1·6 mV per deg C.

Some large-area versions are now appearing with permissible inverse power dissipations of about 50 watts. They will be useful for surge protection and arc-suppression applications.

(5.1.3) **Minority-Carrier Storage.**

In 1950, Meacham and Michaels[52] described a carrier-storage phenomenon in point-contact diodes. This effect is present in all *p–n* junction devices and, in some respects, imposes a limitation of operation. The effect may be described by considering an alloy rectifier using *n*-type material. In such a device, forward current consists mainly of a flow of holes, which are injected into the *n*-type wafer from the fused *p*-type region and will eventually recombine with free electrons either in the wafer or at the base contact. If a rectifier is passing forward current and a reverse voltage is suddenly applied, the holes actually present in the *n*-type wafer are attracted back across the junction before they can recombine and so form a large pulse of reverse current which decreases to the normal value as the holes flow back.

The effect is called 'minority-carrier storage' or, in the case described above, 'hole storage'. In junction diodes and rectifiers it presents a limit to high-frequency operation, especially in fast switching, and for power rectification the decay of such reverse pulses flowing through any series reactance will cause voltage spikes which can be harmful to the rectifier. In *p–n–p* junction transistors hole storage in the base region delays the switching-off of collector current for an appreciable time after the switching-off voltage has been applied between the emitter and base. Thus it will be seen that this effect is very important, and must be carefully considered in the design of *p–n* junction devices.

(5.2) Junction Transistors

Transistors are characterized by the semiconductor material from which they are made, by the method of construction, and by their power and frequency performance. The common feature of all junction transistors is the narrow base region between the emitter and collector regions. The finite transit time of the electrons or holes across this narrow base region sets an upper frequency limit to the transistor, and so the actual thickness of the base region is the first essential design requirement. Table 4 gives theoretical upper frequencies for various base thicknesses, calculated from the diffusion constants of electrons and holes in germanium and silicon. A further requirement for good high-frequency performance is a low value for the product of collector capacitance and series base resistance. This often introduces a limit at a lower frequency than the base-width limitation. Thus, for good high-frequency operation, the junction area must be reduced, and at present this automatically imposes a power limitation on high-frequency transistors. On the other hand, for high-power working the collector and emitter areas must be increased, and so we have a frequency limitation on power transistors.

Table 4

Theoretical Relations between Base Width and Upper Frequency Limit

Base width	Theoretical upper frequency limit			
	Germanium		Silicon	
	p–n–p	*n–p–n*	*p–n–p*	*n–p–n*
in $\times 10^{-3}$	Mc/s	Mc/s	Mc/s	Mc/s
2	0·75	1·5	0·2	0·5
1	3·0	6·0	0·75	2·0
0·5	12·0	24	3·0	7·5
0·2	75	150	20	50
0·1	300	600	75	200
0·05	1200	2400	300	750
0·02	7500	15000	2000	5000

Up to the present, the vast majority of transistors produced have been of the germanium-alloy junction type. This particularly applies to Great Britain and Europe. In the United States, grown junction transistors of both germanium and silicon are also readily available. Grown transistors and surface-barrier transistors are now available to some extent in this country, and more recent types such as diffused transistors, drift transistors, post-alloy diffused transistors and silicon transistors are all making headway and if not already available in pre-production form, will be in the near future.

(5.2.1) **Transistor Characteristic Curves.**

As with thermionic valves, static characteristic curves are produced by the transistor manufacturer to assist the circuit designer. Since with the transistor we have a finite input resistance to the emitter or base, it is necessary to provide, in addition to the collector characteristics, the corresponding emitter characteristics. Also, in view of the possible alternative connections—common base and common emitter—the characteristics will ideally be presented for both arrangements. Fig. 4 shows a set of common-emitter characteristics for a low-power germanium transistor. The similarity of the collector characteristics to those of a pentode valve is readily apparent.

(5.2.2) **Limitations of Transistors and some Comparison with the Thermionic Valve.**

There are some limitations which are fundamentally present to some extent in all types of transistor. The first is the operating temperature. For germanium transistors, the limit of junction temperature is about 100° C, but most manufacturers still specify a limit of 75° C. This allows reasonable operation in ambient temperatures up to 55° C and thus is quite satisfactory for entertainment, and some industrial, applications, at least in this country. With silicon, the limit of junction temperature is in the region of 200–250° C, allowing operation in ambient temperatures up to 175° C. This will be satisfactory for most industrial and military applications, and thus the temperature limitation is not restrictive.

Perhaps a greater limitation is that the transistor parameters vary with temperature. This applies equally to both germanium and silicon. As a result, in most applications it is necessary to include circuits designed to compensate for the effects of these

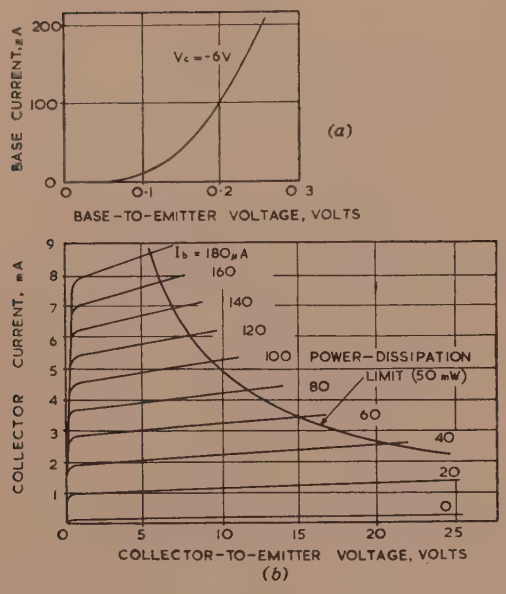


Fig. 4.—Characteristic curves of a typical low-power alloy transistor.

(*a*) Common-emitter input characteristic.
(*b*) Common-emitter collector characteristics.

variations. With germanium transistors operating over a relative small temperature range this is not difficult, but to design these compensating circuits to be effective over a much wider temperature range, in order to take the fullest advantage of silicon transistors, is quite difficult.

The question of the upper frequency limit of operation is one of temporary limitation, depending on the progress of development. Diffused-structure transistors have been made to operate at 1 Gc/s, and doubtless this figure will be increased—the problem being largely one of technology. It should always be remembered, however, that the base width of a 10 Mc/s transistor compares with the grid-cathode spacing of a 1 Gc/s thermionic valve, and so the difficulties must not be underrated.

As to further comparison with the thermionic valve, the obvious advantages of the transistor, such as the absence of a heater, low-voltage operation, small size, rugged construction, etc., have all been stressed. There are, in addition, other advantages. Because of the low junction temperature, the transistor runs relatively cool in operation, and thus associated components are not heated as with the thermionic valve. But perhaps the most far-reaching feature of the transistor is the fact that either electrons or holes can be used as the current carriers, resulting in the alternative *n–p–n* or *p–n–p* types, whereas with the thermionic valve we are limited to a flow of electrons only. This means that a completely new series of transistor circuits can be devised combining the two types, such circuits being impossible with valves.

(5.2.3) Junction-Transistor Types.

In the following Sections, it is intended to mention briefly the basic types of transistor, and in conjunction with Tables 5A and 5B to identify their general features and characteristics.

(5.2.4) General-Purpose Low-Power Transistors.

General-purpose low-power transistors include those with common-base cut-off frequencies around 1 Mc/s, and permissible collector dissipations of about 100 mW. They were, in fact, the first group of junction transistors to become available, and are made either by the grown-junction or alloying process.

(5.2.4.1) *Grown-Junction Types.*[53]

Grown-junction transistors are available in both germanium and silicon, and are invariably of the *n–p–n* type. The grown structure, as described earlier, is cut into bars about 0·020 in square by $\frac{1}{8}$ in long, each bar containing the thin *p* layer. For the general-purpose grown transistor a base width of about 0·001 5 in is used. One of the main problems in manufacture is to make contact to this thin base layer. In practice, a thin wire—gold/gallium for germanium, or aluminium for silicon—is fused to it. Emitter and collector contacts are made to the ends of the bar [Fig. 2(*d*)]. The main disadvantage of the grown transistor is that there is a relatively high internal series resistance between each electrode connection and the actual junction.

Table 5A

TYPICAL CHARACTERISTICS OF BASIC GERMANIUM TRANSISTOR TYPES

Type	Manufacturing process	Structure	Common-base cut-off frequency	Maximum collector power	Maximum collector voltage	Current gain	Availability in United Kingdom
			Mc/s		volts		
General-purpose low power	Alloy	*p–n–p*	1	50/200 mW	10–50	20–100	Yes. Good
	Alloy	*n–p–n*	1	50/200 mW	10–50	20–100	Yes. Restricted
	Grown	*n–p–n*	1	50 mW	20–50	20–100	No
Intermediate power	Alloy	*p–n–p*	0·25	0·5/2 W	30–100	20–50	Yes. Fair
High power ..	Alloy	*p–n–p*	0·1	2/55 W	30–100	20–50	Yes. Fair
Symmetrical switching	Alloy	*p–n–p*	2·5–8·5	70/100 mW	5–20	10–60	Yes. Fair
	Alloy	*n–p–n*	2·5	100 mW	10–20	10–20	Yes. Restricted
High-frequency types	Alloy	*p–n–p*	10	30/100 mW	10–20	20–150	Yes. Good
	Alloy	*n–p–n*	10	30/100 mW	10–20	20–150	Yes. Restricted
	Grown	*n–p–n*	15	50 mW	10–25	20–100	No
	Alloy drift	*p–n–p*	30	50 mW	35	30–100	Yes. Restricted
	Grown tetrode	*n–p–n*	50	50 mW	10–25	20–80	No
	Surface barrier	*p–n–p*	60	20 mW	5–10	10–30	Yes. Fair
	Post-alloy diffused	*p–n–p*	200	50 mW	5–20	10–30	Yes. Preproduction samples
	Diffused	*p–n–p*	500	50 mW	5–20	10–50	No

Table 5B

TYPICAL CHARACTERISTICS OF BASIC SILICON TRANSISTOR TYPES

Type	Manufacturing process	Structure	Common-base cut-off frequency	Maximum collector power	Maximum collector voltage	Current gain	Availability in United Kingdom
			Mc/s		volts		
General-purpose low power	Alloy	*p–n–p*	1	250 mW	30	10–60	Yes. Preproduction samples
	Alloy	*n–p–n*	1	250 mW	30	10–60	Yes. Preproduction samples
	Grown	*n–p–n*	6	150 mW	20	10–90	Yes. Fair
Intermediate power	Grown	*n–p–n*	2	1 W	50–120	10–90	Yes. Restricted
	Diffused	*n–p–n*	2	4 W	60–100	10–40	Yes
High power ..	Diffused	*n–p–n*	1	37·5–80 W	60	10–40	Yes
High frequency	Alloy	*p–n–p*	4	250 mW	30–45	20–80	Yes. Preproduction samples
	Alloy	*n–p–n*	4	250 mW	30–45	20–80	Yes. Preproduction samples
	Grown tetrode	*n–p–n*	15	125 mW	30	20–80	Yes. Fair
	Grown diffused	*n–p–n*	30	125 mW	30	20–90	Yes. Fair
	Diffused	*n–p–n*	100	150 mW	30	25	No

(5.2.4.2) *Alloy-Junction Types.*[54]

The basic structure of the germanium *p–n–p* alloy transistor is shown in Fig. 2(*c*). Two junctions are formed on opposite sides of a thin wafer of *n*-type germanium, which forms the base region, by alloying dots of indium. Collector and emitter diameters are 0·045 and 0·030 in, respectively, and the base width is 0·001 5 in. The maximum collector dissipation is between 100 and 200 mW at 25° C ambient temperature.

In contrast to the grown-junction transistor, the series resistance to all electrodes is relatively low, and thus the structure is very suitable for use in both power and switching transistors.

Production engineering of corresponding *n–p–n* alloy types is in progress, using lead-antimony or indium-antimony combinations alloyed into *p*-type germanium.

As to silicon, a few alloy *p–n–p* types are available, using aluminium alloyed into *n*-type silicon. But owing to the very low minority-carrier lifetime of silicon available at present, a very thin base region is necessary to obtain reasonable current gain. The manufacturing yield to obtain a sufficiently low base width with the alloy process is very low, resulting in restricted output and high cost. Doubtless, the characteristics of the starting silicon will be improved, and then, with a thicker base region, the economics may compare with those of the germanium-alloy transistor.

(5.2.5) **High-Frequency Transistors.**

(5.2.5.1) *Grown-Junction and Alloy-Junction Types.*

As indicated earlier, to extend the high-frequency range of a transistor, it is necessary to decrease the base width, and also both the collector and emitter junction areas. For high-frequency grown-germanium transistors, the cross-section of the bar is reduced to 0·015 in square and the base width to 0·000 5 in, resulting in a maximum frequency of about 15 Mc/s. Similarly the alloy-transistor high-frequency range is extended by reducing the collector and emitter dots to 0·015 and 0·010 in, respectively, and the base width to 0·000 5 in. The selling prices of these two types are quite comparable, and thus the choice depends upon the characteristics preferred for a given application. For high-frequency amplification such as in the r.f. and i.f. stages of radio receivers, the grown-junction type is perhaps preferable because of its low collector capacitance—about 2 to 4 pF, which makes for good stability in tuned-amplifier stages. With the alloy-junction type, the collector capacitance is in the region 10–15 pF, and it is usually necessary to include neutralization to avoid instability due to positive feedback.

(5.2.5.2) *Tetrode Transistor.*

A further development of the high-frequency grown-junction transistor is the so-called tetrode.[55] In this, a second base contact is made opposite to the first. A transverse bias, applied between these two base contacts, has the effect of confining the transistor action to a region near one of them. This reduces the effective area of the transistor, thus decreasing both the collector capacitance and the series base resistance and increasing the maximum operating frequency up to the region 30–40 Mc/s. Silicon tetrodes have recently been developed for use up to 100 Mc/s.

(5.2.5.3) *Alloy Drift Transistors.*

The transistors described previously have a uniform base region, the transit time of the carriers across this region resulting from simple diffusion. In the drift transistor[56] a resistivity gradient is built into the base region so that a field is produced which sweeps the carriers across at a higher speed, thus reducing the transit time. Furthermore, a lower impurity concentration is produced in the base region adjacent to the collector junction, and so the collector capacitance is reduced. Thus an increase in maximum operating frequency is obtained, up to about 30 Mc/s. The resistivity gradient is obtained by diffusing the required impurity into one face of the wafer before alloying the collector and emitter in the usual way, the general structure being similar to the high-frequency alloy transistor.

The alloy drift transistor could well emerge as a preferred standard type, since, to some extent, it combines the attractive features of the high-frequency alloy and tetrode transistors.

(5.2.5.4) *Surface-Barrier Transistors.*

Surface-barrier transistors[57] use a process known as electrolytic jet etching to produce very thin base widths with reduced junction areas [Fig. 2(*e*)]. Tiny jets of a suitable solution are directed on to the opposite surfaces of a germanium wafer to produce structures with emitter diameter of about 0·002 in, collector diameter of about 0·004 in and base width down to 0·000 2 in. A typical operating frequency of the transistors so produced is in the region of 50–70 Mc/s. Owing to the extremely small dimen-

sions, the power rating is low. Silicon types are also being developed and may be available in the near future.

A recent development of this technique is the micro-alloy diffused transistor. A graded base region is formed by diffusion before jet etching and plating, and the emitter and collector electrodes are subsequently alloyed by a short heat cycle.

(5.2.5.5) *Post-Alloy Diffused Transistors.*

Very small base widths have been obtained by a process of diffusion after alloying the emitter region. This process, known as post-alloy diffusion,[16] was pioneered in this country. Indium containing a suitable donor impurity is alloyed into *p*-type germanium, and the structure is then subjected to a high-temperature cycle during which the donor diffuses into the *p*-type wafer to produce a thin *n*-region. This appears to be a very promising method and seems capable of producing units for operation up to 200 Mc/s.

(5.2.5.6) *p–n–i–p Transistors.*[17]

If a thin layer of intrinsically pure semiconductor is included between the collector and the base layer, the collector-to-base capacitance is reduced, resulting in a higher operating frequency. A higher collector voltage will also be possible. Although units of this type were successfully demonstrated some time ago, they have only recently become available commercially.

(5.2.5.7) *Diffused-Junction Transistors.*

The fact that the diffusion process allows good control of penetrations down to a few millionths of an inch indicates that it is an ideal method for producing the thin base layers necessary for very-high-frequency transistors. Such a transistor, capable of operation up to 500 Mc/s, has been described by Lee,[18] and more recent information is that units of this type have operated at well over 1 000 Mc/s. Typical dimensions for such diffused-junction transistors are as follows: base width, 0·000 05 in, emitter, 0·001 in × 0·002 in, and collector 0·005 in diameter. The general structure is shown in Fig. 2(*f*). At least two American companies are producing this type of transistor, and one type is available commercially.

(5.2.5.8) *Spacistor.*

A more recent possibility, still in the development stage is the spacistor.[58] Carriers are injected into the space-charge region of a reverse-biased junction. This makes operation independent of carrier lifetime, and possible operation up to 10 Gc/s has been mentioned.

(5.2.5.9) *Field-Effect Transistor.*[59]

The field-effect transistor is a unipolar device, in which the flow of electrons down a bar of semiconductor material is controlled by means of an inversely biased *p–n* junction on the side of the bar. The depletion layer of the junction extends into the bar, restricting the flow of electrons to an extent depending on the inverse voltage across the junction. A recent version of this device, called the 'technetron', is claimed to operate up to 500 Mc/s.

(5.2.6) **Power Transistors.**

A very comprehensive and complete review of power transistors has recently been presented by Clark.[60] Consideration of the main requirements of increasing junction area and reducing built-in series resistances leads to the conclusion that only two methods of construction are really suitable—alloying and diffusion. Whichever system is used, a thin base width is still necessary to obtain good current gain. In order to keep the series base resistance low, the geometry of the structure will preferably include narrower emitter and base electrodes, with either a multiple strip or concentric circular arrangement.

An essential feature of the construction of power transistors is that a low thermal-resistance connection must be provided from the collector junction to an external surface, which can, in use, be connected to a suitable heat-radiating system. In many applications, the equipment chassis can be used as this heat radiator. The permissible power is dependent not only on the internal thermal resistance between the collector junction and the case, but also on the thermal resistance of the heat-radiating system and the ambient temperature. This whole question becomes somewhat complicated to specify, but manufacturers are now providing extensive details in their data sheets.

A complete range of germanium power transistors is now available with permissible collector dissipations from 0·5 to 55 watts (25 watts in the United Kingdom). These are, in general, *p–n–p* transistors made by the alloy process, but at least one *n–p–n* unit for 7 watts has been announced.

With silicon, a double diffused structure is used. Units are now available up to about 80 watts (37·5 watts in the United Kingdom).

(5.2.7) **Switching Transistors.**

Any junction transistor can be operated as a reasonably efficient electronic switch in the common-emitter arrangement, a large collector current being switched by means of a small base input current.[61] Now appearing on the market are transistors specially designed for this purpose, for application in computer circuits and the like. One popular type is a germanium alloy transistor with a frequency cut-off of several megacycles per second and a collector saturation resistance of about 1 ohm. Some versions use a symmetrical construction, in which the emitter and collector are identical in size. Recent emphasis has been directed to faster switching, and manufacturers are now quoting the rise, storage and decay times of their transistors intended for switching use. For the alloy type previously mentioned, these times are each of the order of a few tenths of a microsecond. Since the collector leakage currents with silicon are much smaller than with germanium, the silicon alloy transistor, when available, will be preferable to the germanium transistor, other parameters being equal.

The so-called avalanche transistor, which operates into the current-multiplication, or avalanche, region of the collector junction characteristic, has properties very suited to switching, with fast rise times. Although many experimental units have been made and demonstrated successfully, no production units are yet available.

Surface-barrier transistors can be used for very fast operation, switching rates in excess of 20 Mc/s being claimed by the manufacturers.

At present, extensive development is proceeding on devices with a bistable characteristic of the thyratron type. The main feature of these devices is that they can be triggered into conduction by means of a small short-duration pulse. Several different devices have been described. Two of these, the unijunction transistor[62] (originally known as the double-base diode) and the 4-layer diode[63] are in production in the United States. Three others are described in References 64, 65 and 66, and samples are available. The controlled rectifier, described in Reference 66, is a 4-layer *p–n–p–n* device with an additional contact made to one of the centre regions for the trigger input. It is initially being developed for use as a controlled power rectifier, but smaller versions will doubtless be developed for low-power circuit use in computer applications. The bistable nature of the characteristics of these devices makes them very attractive for many computing applications, since one of them

an be equivalent to two interconnected conventional transistors. We shall doubtless have considerable development of this type of device in the near future.

(5.2.8) **Transistor Noise Characteristics.**

As with thermionic valves the limit of usefulness of amplification of the transistor is the residual noise produced in the transistor itself. This transistor noise is frequency dependent. Below 1 kc/s it is mainly due to so-called semiconductor noise and increases according to an inverse frequency law. Above 1 kc/s, the noise is relatively constant, and is due to shot and thermal noise. Eventually, as the frequency is increased further, decrease in the transistor gain results in a deterioration of the noise factor. Thus, in order to obtain good noise factors at high frequencies, it is important to use transistors with high cut-off frequency.

It is interesting to compare the low-noise performance of transistors with that of thermionic valves. Typical low-noise a.f. transistors have a minimum noise factor of about 4 dB when operated from a source impedance of 500–1 000 ohms. For such a source impedance, a low-noise thermionic valve has a noise factor of about 10 dB. For higher source impedances, however, the thermionic valve decreases to only 1 or 2 dB, while the transistor increases to about 20 dB. For low-noise operation of transistors, it is important to keep the collector current to a low value, about 0·2 mA.

(5.3) Semiconductor Photo-Electric Devices

The photo-electric properties of semiconductors have been known for a long time, and some development of semiconductor photo-electric devices has taken place over the last ten years. Unfortunately, the commercial demand for photo-electric devices is small compared with that for rectifiers and transistors, which has adversely affected the priority given to their development. However, a small number of junction photo-diodes and photo-transistors are available. In effect, these consist of conventional junction-diode or transistor structures, mounted in such a manner as to receive incident light on the junction. They require a voltage bias supply, and their output current level is increased by increasing the intensity of incident light. A *p–n* junction can also be used as a photo-voltaic device, such as the silicon solar battery,[67] which converts solar energy into electrical energy with an efficiency of up to 11%.

Table 6 gives typical characteristics of the above devices, with the selenium cell for comparison. Characteristics of the cadmium-sulphide cell are also included. Although not within the scope of this review, this is a recent development and is a strong competitor. With sensitivity quoted in milliamperes per lumen, the photo-transistor and cadmium-sulphide cell both show revolutionary improvement over the selenium cell. In making this comparison, however, we must remember that the active area of these new devices is very small, and thus to obtain the full increased sensitivity it would be necessary to include a focusing lens of area equal to the selenium cell. The comparison does, however, indicate the superiority of the new semiconductor photo-electric devices for small-area use, such as punch-card reading in computers, etc.

The solar battery will be of considerable importance if its economics can be improved. Since about 1 m^2 of silicon surface is necessary to give an output of 100 watts in sunlight, and the processed silicon costs at least a few shillings per square centimetre, the cost at present is obviously very high.

The new multi-region devices, such as the 4-layer *p–n–p–n* diode, will lend themselves to photo-electric operation as devices which can be triggered into conduction with a light pulse.

Table 6

TYPICAL CHARACTERISTICS OF VARIOUS PHOTO-ELECTRIC DEVICES

Type	Sensitivity	Wavelength for maximum sensitivity	Applied voltage	Dark currents	Maximum power dissipation	Availability in United Kingdom
	mA/lumen	microns	volts	μA	mW	
Selenium cell	0·3	0·6	—	—	—	Yes. Good
Germanium photo-diode ..	30	1·6	100	50	50	Yes. Fair
Germanium photo-transistor..	300	1·6	25	200	25	Yes. Fair
Silicon solar cell	3	0·8	—	—	—	Yes. Restricted
Cadmium-sulphide cell ..	200	0·5	300	<1	100	Yes. Fair

(6) APPLICATIONS OF SEMICONDUCTOR DEVICES

The rate at which the circuit application of the transistor increased was influenced to some extent by the fact that, when the transistor was invented, the field of electronic circuits based on the thermionic valve had already been in existence for some 30 years. The circuit development engineer had progressed to an advanced understanding of the use of the thermionic valve, and it was reasonable that he would not become really interested in this new device for future products until he was sure that its performance, reliability, availability and cost were all at least comparable with the thermionic valve. Obviously, it would take time for these conditions to occur. Fortunately, quite early on, it became apparent that the junction transistor was ideally suited for use in hearing aids and portable radio receivers, and quite soon, transistorized versions of these items appeared on the market. The requirements of these two applications gave a significant impetus to the production of transistors, resulting in better availability and some reduction in cost. This, in turn, led to more application work—particularly on computer circuits and other industrial applications—and now it would be safe to assume that some transistor application work is proceeding in all electronic circuit development laboratories. As a result, transistor circuit design is now developing into a new and distinct science in its own right, with little relation to the old standard valve circuits.

For semiconductor diodes and rectifiers, the path has been easier. The germanium point-contact diode, and germanium and silicon junction rectifiers, all had, at quite an early stage in their development, a better performance than devices previously in use. Thus, although at first their cost was slightly higher, the improved performance gave rise to sufficient demand for manufacturers to set up large-scale production, which soon resulted in good availability at attractive selling prices.

The increase of the circuit application of transistors is illustrated in Fig. 5, which shows the number of technical articles on

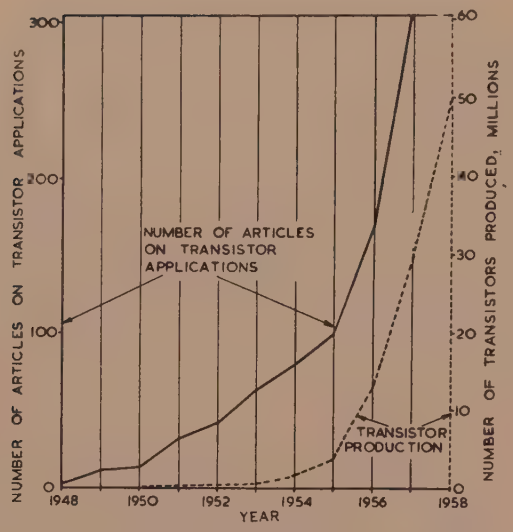


Fig. 5.—Build-up of the application of transistors.

such applications published each year from 1948 onwards. It is interesting to compare this with the increase in the production of transistors, which is given for comparison.

(6.1) Application of Semiconductor Diodes and Rectifiers

The two major applications of germanium point-contact diodes are in television receivers and computers. In the former, they are used for detection, limiting, d.c. restoration and synchronizing impulse separation; in the latter, for various pulse-manipulation circuits such as clamping, limiting and steering. In both cases, economics are very important—a typical industrial computer may use up to 10000 diodes. Thus the relatively inexpensive point-contact type of diode will find application for a long time to come—at least until the miniature junction type can be manufactured as cheaply. A typical flip-flop circuit which illustrates the possible use of large numbers of diodes in computer circuits is shown in Fig. 6.

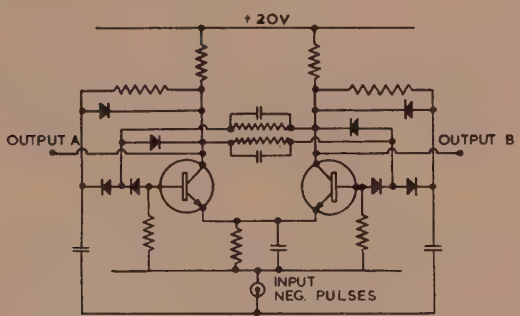


Fig. 6.—Typical transistor flip-flop circuit illustrating the wide use of diodes.

The application of semiconductor junction rectifiers is mainly, of course, to the rectification of power supplies, and they have already replaced earlier types of rectifier in many instances. Two points must be noted in their application. First, as their forward resistance is so low, the circuit must be designed to limit the peak surge current, usually by including a suitable series resistance, and secondly, it is necessary to protect the rectifiers from short-duration over-voltage spikes.

In the high-power field, germanium and silicon rectifiers are finding extensive application for electrolysis, electroplating, traction and many other industrial uses.

(6.2) Basic Transistor Circuit Design

There are certain characteristics of the transistor which necessitate differences in circuit arrangement from the customary thermionic-valve circuits. These are as follows:

(*a*) Very low input impedance.
(*b*) Temperature-dependent bias setting.
(*c*) Variation of current gain with emitter current.
(*d*) Fundamental internal feedback.

The circuits evolved to take care of these effects, together with established basic electronic circuits such as interstage coupling circuits and new arrangements made possible by combinations of *p–n–p* and *n–p–n* types, all comprise the new transistor circuit design.

For instance, to take care of the temperature dependence of bias setting, various bias-stabilization circuits have been evolved. These, in effect, apply negative feedback to the d.c. bias circuit. To reduce the effect of the variation of gain with emitter current, negative feedback is applied to the a.c. signal in much the same way as with valve circuits. The presence of internal feedback in the transistor means that the input impedance is dependent on the load resistance and the output impedance is dependent on the source resistance. At high frequencies, the internal feedback via the collector capacitance to the base resistance can result in positive feedback. If this is excessive, neutralization is necessary.

For pulse work, advantage is taken of the ideal switching characteristics of transistors, the low collector saturation voltage, and combinations of *p–n–p* and *n–p–n* types to produce simple circuits which have no direct thermionic valve counterpart.

(6.3) Transistor Applications

Finally, mention is made of a few typical fields of application, to illustrate what has been or is being carried out with transistors. A general paper on transistor circuits and applications was read before The Institution in 1957 by Milnes.[68]

(6.3.1) Hearing Aids.

Transistors have now completely replaced thermionic valves in hearing aids. Typical transistor hearing aids operate from one small 1·5-volt pen-torch cell, and can be made small enough to be worn behind the ear or in spectacle arms.

(6.3.2) Radio Receivers.

The application of the transistor to radio receivers has, until now, been confined to the portable and car types. Here, again, the attractive feature is the low battery consumption. The author has a pocket transistor radio purchased in January, 1956. It is powered by four 1·5-volt pen-torch cells, and although it has been in steady intermittent use throughout these three years, it is still operating satisfactorily on the original batteries. With car radio, operation direct from the car battery effects economy by avoiding the use of the vibrator power supply.

(6.3.3) A.F. Amplifiers.

Low-distortion a.f. amplifiers for outputs of up to 20 watts have been described[69] and are finding application for public-address systems. The output power will undoubtedly be extended for cinema amplifiers and the like.

(6.3.4) Television Receivers.

There is little incentive to use transistors in television receivers from the point of view of saving space or reducing power consumption, but there is a definite incentive to use them if it helps to reduce the receiver cost. This represents real possibilities. For instance, the use of power transistors in magnetic scanning circuits could well result in a cheaper overall system. A com-

plete television receiver using transistors throughout has recently been described by Overton.[71]

(6.3.5) **Telephone Systems.**

There are obvious applications for the transistor in many component units of telephone systems, such as line amplifiers, repeaters, telephone switching, signalling, dialling, ringing and carrier equipment. Several field trials are under way by the Post Office, in typical installations, and transistors will gradually be introduced into the existing telephone system as their suitability and reliability are proved.

In parallel, extensive development is under way towards the all-electronic telephone exchange, in which many circuits will be built around the transistor and the semiconductor diode.

(6.3.6) **Computers.**

The various features of the transistor are all used to considerable advantage in computer applications. In an equipment which can use as many as 10 000 transistors, the small size is an obvious attraction. The low operating power, coupled with absence of heater supply, also contributes to small equipment size and keeps the internal heat dissipation low. A typical computer[70] uses 5 000 transistors and needs a d.c. power supply of only 20 watts. But perhaps the most attractive feature of the transistor in computers is the anticipated reliability—the above computer has so far given 15 million transistor hours of operation without a single failure.

Probably of interest to all is the fact that transistors are used in the counting circuits of Ernie—the computer that picks out the Premium Bond winners.

(6.3.7) **Industrial Applications.**

The application of transistors to industry is only just now beginning to gain momentum. Systems have been described using transistors in servo-amplifiers for machine-tool control, and as control elements for motor speed control. Transistor d.c./a.c. convertors have been developed for output powers up to about 1 kW. A typical medium-power convertor is a unit to provide fluorescent lighting in locomotive coaches from the 24-volt d.c. supply. In an experimental convertor for this purpose, two power transistors supply an output of 40 watts to the fluorescent lamp with an overall efficiency of 65%. This type of convertor can be arranged to give a 3-phase a.c. output if so desired, by the use of a 3-phase oscillator and transformer.

The use of power transistors and controlled rectifiers as power switches is being investigated with a view to replacing contactors in electrical control gear. Initial results show considerable promise.

(6.3.8) **Instrumentation.**

Many transistorized instruments have been developed and some are already commercially available. Typical of these are a.c. microammeters, a.c. millivoltmeters, frequency meters, radiation detectors and monitors, strain gauges, metal detectors, and portable oscilloscopes. A typical experimental transistor oscilloscope uses 30 transistors and weighs, complete with internal batteries, 15 lb. Also important is the use of transistors in radiosonde equipment and earth satellites, where their small size, weight and power requirements are ideal.

(6.3.9) **Applications to Military Equipment.**

The possible uses of transistors in military equipment are almost boundless—line and radiocommunication, radar of all types, aircraft equipment, guided missiles and many other specialist applications too numerous to mention. Silicon devices with their extended temperature range will undoubtedly find extensive application.

(6.4) Transistor Reliability

In 1952, a life of 70 000 hours was forecast for the point-contact transistor, based on the extrapolation of overload operation. Subsequent experience has indicated that this is a very reasonable estimate. With junction transistors, manufacturers are now quoting estimated lives of 100 000 hours. A few typical examples of operating reliability can be quoted. With one particular type of hearing aid, 70 000 units of which are in the field, only one known failure occurred in two years of operation. As mentioned previously, 5 000 transistors in the Leprechaun[70] computer have operated for 3 000 hours without a single failure.

(7) FUTURE TRENDS

The transistor art has been continuously changing over the last few years, and no doubt will continue to do so for some time in the future. While the processes of diffusion and alloying appear to be possible permanent features of device fabrication, it would be fairly safe to predict that, in the search for higher-frequency and higher-power operation, new structures and techniques will be devised that may have a pronounced effect on the redesign of existing devices. Furthermore, devices such as the field-effect transistor, Shockley's unipolar analogue transistor, the spacistor, new multi-region devices, parametric amplifiers, Hall-effect devices and thermo-electric devices will each be the subject of considerable development, and some of them will doubtless come to fruition.

It is confidently expected that the high-frequency limit of the diffused germanium transistor will be increased to several thousand megacycles per second, and there seems to be no fundamental reason why power transistors with collector dissipations approaching a kilowatt should not be possible. Transistors for high-speed switching, up to one or two hundred megacycles per second, will become available, and a range of *p–n–p–n* devices for computer and power-control applications will become established.

In addition to germanium and silicon there are a number of possible starting materials, known as compound semiconductors, and much effort is going over to them for special applications. For example, work is proceeding on silicon carbide and indium phosphide, which, because of their high energy gap, should be suitable for high-temperature devices—the former up to 500° C. Compounds with high mobilities, such as indium antimonide and indium arsenide, are being developed for use in Hall-effect and photo-electric devices. At present, however, there is still much to accomplish, both in the preparation of these compounds and also in their application to devices.

(8) CONCLUSIONS

In the brief space of eight years, the junction transistor has reached the state, where, in many applications, it is competitive both in performance and cost with the thermionic valve, which has had over 50 years of development. At the same time, the semiconductor rectifier has overtaken previously established rectifiers in both the low- and high-power fields. From time to time, the question has been raised 'Will transistors oust thermionic valves', and opinion has swayed to and fro as enthusiasm has waxed and waned. The author is one of those who feels quite strongly that solid-state devices will be developed to perform all the functions of existing thermionic valves, and may eventually almost completely replace them.

(9) ACKNOWLEDGMENT

The author wishes to acknowledge the co-operation of the many manufacturers who have provided him with various details of their semiconductor devices.

(10) REFERENCES

(1) BLEANEY, B., RYDE, J. W., and KINMAN, T. H.: 'Crystal Valves', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 847.

(2) DOUGLAS, R. W., and JAMES, E. G.: 'Crystal Diodes', *Proceedings I.E.E.*, Paper No. 1001 R, March, 1950 (**98**, Part III, p. 157).

(3) SCOTT, T. R.: 'Crystal Triodes', *ibid.*, Paper No. 990 R, May, 1950 (**98**, Part III, p. 169).

(4) SCOTT, T. R.: 'Crystal Valves', *Journal of the Television Society*, 1958, **8**, p. 401.

(5) 'Semiconductor Eelectronics' (Semiconductor Information Service, Cambridge, Mass., U.S.A.).

(6) POWELL, A. R., LEVER, F. M., and WALPOLE, R. E.: 'The Extraction and Refining of Germanium and Gallium', *Journal of Applied Chemistry*, 1951, **1**, p. 541.

(7) PFANN, W. G.: 'Principles of Zone Melting', *Journal of Metals*, 1952, **4**, p. 747.

(8) KECK, P. H., and GOLAY, M. J. E.: 'Crystallisation of Silicon from a Floating Liquid Zone', *Physical Review*, 1953, **98**, p. 1297.

(9) TEAL, G. K., SPARKS, M., and BUEHLER, E.: 'Growth of Germanium Single Crystals containing *p–n* Junctions', *ibid.*, 1951, **81**, p. 637.

(10) WILSON, J. M.: 'The Chemical Purification of Germanium and Silicon', *Progress in Semiconductors*, 1958, **3**, p. 27.

(11) DACEY, G. C., and THURMOND, C. D.: '*p–n* Junctions in Silicon and Germanium. Principles, Metallurgy and Applications', *Metallurgical Review*, 1957, **2**, p. 157.

(12) PFANN, W. G.: 'Redistribution of Solutes by Formation and Solidification of a Molten Zone', *Journal of Metals*, 1954, **6**, p. 294.

(13) HALL, R. N., and DUNLAP, W. C.: '*p–n* Junctions Prepared by Impurity Diffusion', *Physical Review*, 1950, **80**, p. 467.

(14) FULLER, C. S., and DITZENBERGER, J. A.: 'Diffusion of Donor and Acceptor Elements in Silicon', *Journal of Applied Physics*, 1956, **27**, p. 544.

(15) CORNELISON, B., and ADCOCK, W. A.: 'Transistors by the Grown-Diffused Technique', I.R.E. Wescon Convention Record, 1957, Part 3, p. 22.

(16) BEALE, J. R. A.: 'Alloy-Diffusion: A Process for Making Diffused Base Junction Transistors', *Proceedings of the Physical Society*, B, 1957, **70**, p. 1087.

(17) EARLY, J. M.: '*p–n–i–p* and *n–p–i–n* Junction Transistor Triodes', *Bell System Technical Journal*, 1954, **33**, p. 517.

(18) LEE, C. A.: 'A High Frequency Diffused Base Germanium Transistor', *ibid.*, 1956, **35**, p. 23.

(19) ANDERSON, D. L.: 'Adhesion of Solids—Principles and Applications', *Bell Laboratories Record*, 1957, **35**, p. 441.

(20) BAKER, D., and YEMM, H.: 'The Depth of Surface Damage Produced by Lapping Germanium Monocrystals', *British Journal of Applied Physics*, 1957, **8**, p. 302.

(21) BIONDI, F. J.: 'Cleaning Electron Device Parts', *Bell Laboratories Record*', 1958, **36**, p. 288.

(22) LITTLE, J. B. (ED. HUNTER, L. P.): 'Handbook of Semiconductor Electronics' (McGraw-Hill), p. 9·1.

(23) SHOCKLEY, W.: 'The Theory of *p–n* Junctions in Semiconductors and *p–n* Junction Transistors', *Bell System Technical Journal*, 1949, **28**, p. 435.

(24) PRITCHARD, R. L.: 'Advances in the Understanding of the *P–N* Junction Triode', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 1130.

(25) HALL, R. N.: 'Power Rectifiers and Transistors', *ibid.*, 1952, **40**, p. 1512.

(26) SABY, J. S.: 'Junction Rectifier Theory', *Proceedings of the Physical Society*, Report on the Rugby Conference, 1956, p. 39.

(27) KINMAN, T. H., CARRICK, G. A., HIBBERD, R. G., and BLUNDELL, A. J.: 'Germanium and Silicon Power Rectifiers', *Proceedings I.E.E.*, Paper No. 1936 U, October, 1955 (**103** A, p. 89).

(28) MCKAY, K. G.: 'Avalanche Breakdown in Silicon', *Physical Review*, 1954, **94**, p. 877.

(29) MILLER, S. L.: 'Avalanche Breakdown in Germanium', *ibid.*, 1955, **99**, p. 1234.

(30) EARLY, J. M.: 'Effects of Space-Charge Layer Widening in Junction Transistors', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 1401.

(31) EARLY, J. M.: 'Design Theory of Junction Transistors', *Bell System Technical Journal*, 1953, **32**, p. 1271.

(32) WEBSTER, W. M.: 'On the Variation of Junction Transistor Current Amplification Factor with Emitter Current, *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 914.

(33) PRITCHARD, R. L.: 'Frequency Variations of Current Amplification Factor for Junction Transistors', *ibid.*, 1952, **40**, p. 1476.

(34) PRITCHARD, R. L.: 'Frequency Variations of Junction Transistor Parameters', *ibid.*, 1954, **42**, p. 787.

(35) KROEMER, H.: 'The Drift Transistor', *Naturwiss*, 1953, **40**, p. 578.

(36) BARRER, R. M.: 'Diffusion in and through Solids' (Cambridge University Press, 1951).

(37) DUNLAP, W. C.: 'Diffusion of Impurities in Germanium', *Physical Review*, 1954, **94**, p. 1531.

(38) EBERS, J. J., and MOLL, J. L.: 'Large Signal Behaviour of Junction Transistors', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 1761.

(39) FLETCHER, N. H.: 'Some Aspects of the Design of Power Transistors', *ibid.*, 1955, **43**, p. 551.

(40) KIDD, M. C., HASENBERG, W., and WEBSTER, W. M.: 'Delayed Collector Conduction, a New Effect in Junction Transistors', *RCA Review*, 1955, **16**, p. 16.

(41) MOLL, J. L., TANENBAUM, M., GOLDEY, J. M., and HOLONYAK, N.: '*p–n–p–n* Transistor Switches', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 1174.

(42) ALDRICH, R. W., and HOLONYAK, N.: 'Multiterminal *p–n–p–n* Switches', *ibid.* 1958, **46**, p. 1236.

(43) VAN DER ZIEL, A.: 'Noise in Junction Transistors', *ibid.*, 1958, **46**, p. 1019.

(44) PLUMMER, A. R. F.: 'Observations on the Growth of Excess Current in Germanium *p–n* Junctions', *Proceedings of the Physical Society*, B, 1956, **69**, p. 539.

(45) ERIKSEN, W. T., STATZ, H., and DE MARS, G. A.: 'Excess Surface Currents on Germanium and Silicon Diodes', *Journal of Applied Physics*, 1957, **28**, p. 133.

(46) KINGSTON, R. H.: 'Review of Germanium Surface Phenomena', *ibid.*, 1956, **27**, p. 101.

(47) *Proceedings I.E.E.*, 1958, **105** B, Supplement No. 11.

(48) SCOTT, T. R.: 'Transistors and Other Crystal Valves' (Macdonald and Evans).

(49) PFAFFENBERGER, W.: 'Technology of Silicon Rectifiers', *Siemens Zeitschrift*, 1958, **32**, p. 115.

(50) FORSTER, J. H., and ZUK, P.: 'Millimicrosecond Diffused Silicon Computer Diodes', I.R.E. Wescon Convention Record, 1958, Part 3, p. 122.

(51) SMITH, D. H.: 'Suitability of the Silicon Alloy Junction Diode as a reference Standard in Regulated Metallic Rectifier Circuits', *Electrical Engineering*, 1955, **74**, p. 43.

(52) MEACHAM, L. A., and MICHAELS, S. E.: 'Observations of the Rapid Withdrawal of Stored Holes from Germanium Transistors and Varistors', *Physical Review*, 1950, **78**, p. 175.

(53) SHOCKLEY, W., SPARKS, M., and TEAL, G. K.: '*p–n* Junction Transistors', *ibid.*, 1951, **83**, p. 157.

(54) SABY, J. S.: 'Fused Impurity *p–n–p* Junction Transistors', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 1358.

(55) WALLACE, R. L., SCHIMPF, L. G., and DICKTEN, E.: 'A Junction Transistor Tetrode for High Frequency Use', *ibid.*, 1952, **40**, p. 1395.

(56) KESTENBAUM, A. L., and DITRICK, N. H.: 'Design, Construction and H.F. Performance of Drift Transistors', *RCA Review*, 1957, **18**, p. 12.

(57) BRADLEY, W. E.: 'Principles of the Surface Barrier Transistor', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 1702.

(58) STATZ, H., and PUCEL, R. A.: 'The Spacistor—A New Class of High Frequency Semiconductor Devices', *ibid.*, 1957, **45**, p. 317.

(59) DACEY, G. C., and ROSS, I. M.: 'The Field Effect Transistor', *Bell System Technical Journal*, 1955, **34**, p. 1149.

(60) CLARK, M. A.: 'Power Transistors', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 1185.

(61) BRIGHT, R. L.: 'Junction Transistors used as Switches', *Transactions of the American I.E.E.*, 1955, **74**, Part I, p. 111.

(62) LESK, I. A., and MATHIS, V. P.: 'The Double Base Diode—A New Semiconductor Device', *The Institute of Radio Engineers Convention Record*, 1953, Part 6, p. 2.

(63) SHOCKLEY, W.: 'The Four Layer Diode', *Electronic Industries*, 1957.

(64) MUELLER, C. W., and HILIBRAND, J.: 'The Thyristor—A New High Speed Switching Transistor', *The Institute of Radio Engineers Transactions on Electronic Devices*, 1958, ED–5, p. 2.

(65) KRUPER, A. P.: 'The Dynistor Diode—A New Device for Power Control' (Communication from the Westinghouse Corporation).

(66) FRENZEL, R. P., and GUTZWILLER, F. W.: 'Solid State Thyratron Switches Kilowatts', *Electronics*, 1958, **31**, p. 52.

(67) PRINCE, M. B.: 'Silicon Solar Energy Converter', *Journal of Applied Physics*, 1955, **26**, p. 534.

(68) MILNES, A. G.: 'Transistor Circuits and Applications', *Proceedings I.E.E.*, Paper No. 2368 R, May, 1957 (**104** B, p. 565).

(69) JONES, D. D., and HILBOURNE, R. A.: 'Transistor A.F. Amplifiers' (Iliffe).

(70) 'The Leprechaun Computer', *Computers and Automation*, 1957, **6**, p. 10.

(71) OVERTON, B. R.: 'Transistors in Television Receivers', *Journal of the Television Society*, 1958, **8**, p. 444.

621.394.64 : 621.395.64 : 621.382

The Institution of Electrical Engineers
Paper No. 2722 R
Sept. 1958
©

THE APPLICATION OF TRANSISTORS TO LINE COMMUNICATION EQUIPMENT

By H. T. PRIOR, B.Sc., D. J. R. CHAPMAN, B.Sc., and A. A. M. WHITEHEAD, B.Sc., Associate Members.

(*The paper was first received* 15*th March, and in revised form* 28*th June*, 1958. *It was published in September*, 1958, *and was read before the* RADIO AND TELECOMMUNICATION SECTION 5*th January, the* MERSEY AND NORTH WALES CENTRE 19*th January, the* SOUTH-EAST SCOTLAND SUB-CENTRE 17*th March, the* SOUTHERN CENTRE 25*th March, and the* SOUTH-WEST SCOTLAND SUB-CENTRE 14*th April*, 1959.)

SUMMARY

The paper discusses the ways in which the use of transistors affects the design of communication transmission equipment for land lines. Transistors have many advantages over valves, but in order to exploit these advantages in practical equipment it has been necessary to wait for transistors of sufficient reliability.

The first step is to design equipment which is compatible with existing valve equipment. In doing this, it is practicable with transistors to be much more free in the use of active circuits, and this makes possible new approaches to the design problems.

A more important aspect is that with transistors new types of system become practicable. Battery operation of relatively complex equipment at unattended stations becomes possible. Very small, low-power repeaters needing no buildings alter the balance between repeater cost and cable cost in the direction of cheaper cables with more frequent repeaters. Freedom in the use of active elements facilitates the use of the digital circuits needed for pulse systems.

The prolific development of semiconductors makes it difficult to steer a reasonable course between making use of progress and restricting the number of types in use. It is likely that eventually the only active elements used will be semiconductor devices.

(1) BACKGROUND

(1.1) Introduction

Long-distance land-line communication has been made possible on an economic scale only by the use of active elements to amplify the signals and counteract the attenuation of the line. Multiplex techniques[1] have permitted the simultaneous transmission of many channels on a single pair of wires, but at the expense of a greater number of active elements to cover the increased attenuation at higher frequencies and the losses of the terminal multiplexing and separating equipment. The thermionic valve is well established as the active element in present-day equipment, and in considering the potentialities of transistors in such applications it is revealing to compare the relative merits of transistors and valves. This comparison has been made many times before,[2–4] but it is necessary here to emphasize those aspects of particular importance in line equipment. The type considered here is the germanium fused-alloy-junction transistor,[5] since this has been found to combine relative ease of manufacture with good reliability in service and good general performance. Other types worthy of consideration are dealt with in Section 3.

The principal advantages and disadvantages of transistors, in order of importance, are as follows:

Advantages	*Disadvantages*
Power efficiency.	Temperature sensitivity.
Life.	Extra significant parameters.
Power range flexibility.	Voltage and power limitations.
Linearity.	Frequency limitations.
No heater.	Noise.
Mechanical robustness.	
Size.	

For use in line-communication equipment the most outstanding advantages of transistors over thermionic valves are the lower power consumption and the greater life expectancy; both stem largely from the solid-state construction and the avoidance of a high-temperature emitter in the form of a heated cathode. This gives an immediate saving in heater power, and furthermore, since the characteristics of a transistor permit satisfactory operation at very low bias voltage and current, there is a considerable saving in supply power for stages handling low signal levels, while for high-level stages the degree of linearity is such that the supply power is still of the same order as the maximum signal to be delivered. This flexibility in catering for a wide range of power levels makes it a practical proposition to design a great variety of equipment using only one type of transistor for all purposes. It is considered that the simplification in the initial stocking and the ensuring of continuity in availability of transistors which is possible with the use of one type, or at least only a few types of transistor, is worth any slight circuit complication; for instance, the use of a push-pull output stage may be justified at the higher power levels to be dealt with, rather than a single transistor of higher power rating.

The greater degree of linearity in transistors is of particular importance in frequency-division-multiplex equipment, in which non-linear distortion causes inter-channel crosstalk.[6] This distortion is reduced to acceptable levels by the application of negative feedback: it is found that less feedback need be applied to a transistor amplifier than to a valve amplifier.

The solid construction of the active region together with the absence of a heater filament enable the transistor to be compact and robust. Although line-communication equipment is not normally expected to be subject to severe mechanical shocks or vibrations while in service, there are advantages in fitting the transistors in their final positions during assembly. The equipment can then be tested and shipped intact, and installation is simpler and speedier. The size and shape of the transistors in general use are typical of the other components associated with them, and a neat and compact assembly is possible. The low power consumption removes the restrictions on the close spacing of temperature-sensitive components which may be necessary for a hotter-surfaced device. Also, in consequence of its greater reliability, ease of replacement is not of such great importance, so it is not necessary for the transistor to be mounted only on the most accessible surfaces of the equipment; this allows a closer physical integration with its associated circuit.

It is unfortunate that by its nature a semiconductor device is temperature-sensitive. In alloy-junction transistors the effects of primary importance with high-temperature operation are an increase in the collector standing current and an increase in the current gain.[7] These effects are of a temporary nature for moderate temperatures, but permanent damage may result at very high temperatures, consisting of a decrease in current gain and a large increase in collector conductance. Considerable advances have been made in the design of transistors to safeguard them against high-temperature damage, first by eliminating contaminants and secondly by improving the heat conduction from the junction to permit high dissipation without the occurrence of an excessive junction temperature.[8, 9] In this respect

The authors are with Standard Telephones and Cables, Ltd.

some compromise is necessary between transistor size and power handling. Long transistor life can be expected only if the circuits are designed to limit the dissipation to the maximum rating which applies for the expected range of ambient temperatures, with the particular cooling methods employed. As a corollary, it is important to avoid the use of mixed valve and transistor equipment. The provision of a simple heat sink will help in conducting the heat away from the transistor case. The effects of the junction-temperature changes which do occur are minimized by suitable biasing methods which reduce collector standing-current changes,[10, 11] and by feedback to reduce changes in gain.

Silicon transistors have a much smaller change of collector current with temperature under typical ambient conditions, owing to the much smaller saturation-current component,[12] but they still possess an appreciable temperature-dependence of current gain. The silicon grown-junction transistor is more troublesome in this respect than the alloy junction.[13]

Even at low frequencies transistor circuit design is complicated by there being a greater number of significant parameters, some of them being dependent on operating conditions.[14] This difficulty, however, can be surmounted by suitable design techniques and, where parameters controlling impedances are likely to vary, by the use of conservative types of circuit and the application of feedback.

A typical general-purpose transistor in use to-day has a maximum power-handling capacity for class-A operation of about 50 mW, which is ten times lower than that of the most common valves it is expected to replace, while the maximum voltage rating is also reduced by a similar factor. These limitations are to some extent offset by the much wider current range, which permits the maximum power to be delivered to a wide range of load impedances. Operating voltages are such that it is convenient to supply equipment from 24–50-volt supplies, which are often already available in exchange buildings. The special precautions necessary when high voltages are present are no longer needed.

The compactness of a transistor element with its low thermal capacity introduces a danger of damage in soldering the leads, or those of an adjacent component, while the relatively low turnover voltage makes it important to prevent the occurrence of even low-energy voltage surges. However, the same sort of breakdown phenomenon has led to the introduction of compact semiconductor protecting devices—Zener diodes[15] and Varistors[16]—which can suitably limit these surge voltages at vital points.

Transistors have already done much to live down their early reputation for high noise generation, and for present applications in land-line transmission equipment a noise figure of 10 dB above thermal noise may be assumed. The greater noise which occurs at low frequencies could affect certain applications such as music circuits, and special care is required in catering for their needs.[17] At high frequencies the decrease in signal/noise ratio occurs only at frequencies at which feedback is decreasing and satisfactory operation is no longer obtainable. No trouble is encountered with noise due to microphony.

A particular limitation in high-frequency equipment is the decrease in current gain with frequency; this is not a reactance effect, but follows from the dispersion of the minority carriers in traversing the base region of the transistor.[18] For existing production alloy-junction transistors the cut-off frequency for common-base operation is in the range 1–20 Mc/s. Unfortunately, with existing manufacturing techniques the higher-frequency values are obtained only at the expense of lower voltage and power ratings, and a suitable compromise must be made in specifying requirements.

(1.2) Transistors as a Practical Proposition

To assess whether we are justified at present in the adoption of transistors in communication equipment it is necessary to be more specific regarding the conditions of operation and the properties of the transistor type considered.

Dealing first with low-frequency and moderate carrier-frequency applications up to about 200 kc/s, we may consider a transistor having a common-base cut-off frequency of 1 Mc/s and a rated dissipation of 50 mW at 50° C. The common-emitter current gain is typically 50, with a range of 35–70. The maximum collector voltage is 30 volts and the current gain remains steady for collector bias currents up to some 25 mA. A typical noise figure is 5 dB at 100 kc/s.

It is considered that the highest ambient temperature in which the transistor equipment will be expected to operate for any length of time is 45° C, and allowing 5° C rise of temperature from room ambient to bay temperature (assuming a completely transistorized equipment) the maximum permitted transistor dissipation will be 50 mW: this figure can be extended if a heat sink is provided. These dissipation figures are conservative and could no doubt be exceeded for short periods, but they have been kept low to ensure the utmost reliability in service.

For audio-frequency operation such a transistor will give a stage gain of about 32 dB in capacitor-coupled common-emitter stages, or about 24 dB in transformer-coupled common-base stages. The former arrangement will require feedback to reduce the 5 dB range in stage gain due to transistor spread; the latter would vary by only 0·1 dB per stage for the range of transistors. In either case, 6 dB would be lost in feedback to match the output to the load, and probably some padding loss or additional feedback to ensure a good input match.

Thus a 3-stage transistor amplifier can supply a highly stable gain of some 55 dB with closely matching terminal impedances. A corresponding result but with higher output power could be obtained by a 2-stage amplifier using CV138 valves. For comparable output, the final stage of the transistor amplifier would require several transistors.

For higher-carrier-frequency applications, up to about 1 Mc/s, another transistor type with a cut-off frequency of about 10 Mc/s is required. With alloy-junction construction the voltage and power ratings are each reduced by about 50% in extending the frequency range, but with suitable circuit design these limitations need not be serious. It is convenient that the lower impedance levels occurring in wide-band circuits lead to reduced voltage swings for given power levels. Wide-band stage gains of about 15 dB may be realized, compared with about 26 dB for a valve stage.

Taking a system as a whole, the ratio of transistors to valves for the same duty is about 2 : 1. However, with its triode construction, the transistor tends to require fewer components per stage than the valve.

It is found that there is little difference in the relative performance of common-emitter and common-base multi-stage amplifiers at audio frequencies so far as gain stability, frequency response and distortion are concerned. The common-emitter arrangement is in many ways more flexible, but at the same time more demanding in the complexity of the feedback network to ensure operation with the widest range of transistors.

(1.3) Developments in Other Components

The introduction of transistors gives a new importance and new impetus to the reduction in size and cost of resistors, inductors and capacitors—a trend which has been in evidence for many years.[19, 20]

In components which are parts of active circuits the low

voltages and powers occurring with transistors have a direct effect on size and cost. Small low-voltage capacitors of the metallized-paper,[21] coated-foil[22, 23] and tantalum-electrolytic[24] types may be used to replace paper capacitors. The size and cost of inductors can also be reduced, since the d.c. power required for a given signal power and linearity is much smaller with transistors than with valves. The lower voltages and temperatures will improve the component reliability towards that of passive circuits, thus further reducing maintenance.

It is important that there should also be progress in components for passive circuits or the improvements in those for active circuits will yield steadily diminishing returns. This progress is being achieved, the most important features being the steady improvements in ferrite material for inductors[25] and in capacitors using plastic dielectrics.

(2) PRESENT SITUATION

(2.1) Transistor Versions of Existing Types of System

(2.1.1) Introduction.

The first step in applying transistors to communication equipment is to produce systems which, in their main features, are similar to established types of valve system. This leads to arrangements which are economically attractive compared with their valve prototypes and compatible with existing valve systems. There are two important aspects of compatibility—the inter-working of transmitting and receiving terminals and the use of a multiplicity of wires following the same route to carry a multiplicity of systems. The basic group of twelve carrier channels in the range 60–108 kc/s established by the C.C.I.T.T.[1] is a case where inter-working between transmitting and receiving terminals of different types is important. Open-wire and multi-conductor cable systems are cases where systems using the same route must be very similar in order to reduce crosstalk between them.

The advantages of using transistors in such cases may be summarized as follows:

(*a*) Power consumption is reduced and power supply simplified, since a single d.c. supply suffices. This is particularly advantageous if a 24- or 50-volt station battery is already installed.
(*b*) Greater freedom in the use of active elements permits economies in the passive elements.
(*c*) Because of longer life, maintenance is reduced.
(*d*) Equipment construction is simplified by the lower heat dissipation. The smaller amount of heat also makes it possible to have more systems in a given station for the same temperature rise.

These points are illustrated in Sections 2.1.2–2.1.7 by reference to equipment which covers a wide field and with which the authors have had practical experience. The equipment exploits the ability of transistors, discussed in Section 1, to work efficiently over a wide range of circuit conditions. Only one type is used, having a dissipation of 50 mW and a cut-off frequency of 1 Mc/s. (No valves are used.) The saving in power consumption of the equipment is listed in Table 1.

Table 1

COMPARISON OF VALVE AND TRANSISTOR EQUIPMENT

Equipment	Power consumption		Relative volume of transistor version
	With valves	With transistors	
	watts	watts	%
Telegraph regenerative repeater	10	1	40
Audio repeaters ..	5	0·5	80
V.F. telegraph channel panel (a.m.)	7·5	1	80
V.F. telegraph channel panel (f.m.)	12·5	1	30
Telephone channel panel	7	1	80
3-channel open-wire terminal	270	33	33
12-channel system repeater	50	5	25

(2.1.2) Telegraph Regenerative Repeater.

The function of a regenerative repeater[26] is to accept a distorted telegraph signal, the distortion being essentially displacement of the signal instants from their correct relative times (although waveform distortion may also occur) and to retime the signals and re-form the waveform. Then, provided that the original time distortion is not greater than half an element, the output signal has no telegraph distortion. A secondary function of regenerative repeaters, particularly applicable to the start-stop type of transmission in very wide use, is to make certain decisions about how a distorted telegraph character shall be retransmitted beyond those needed for mere retiming of the signal instants. The following features, for example, may be required when regenerating 50-baud start-stop signals in which each character consists of a start element, five message elements and one (or $1\frac{1}{2}$) stop elements:

(*a*) Rejection of start signals of less than some specified minimum duration.
(*b*) Addition of a stop signal to a character received without one.
(*c*) Accurate regeneration of a 'space' condition longer than six elements ('space' being the condition used for the start signal).

From this brief account it may be seen that a start-stop regenerative repeater may be regarded as a small computer, both the retiming function and the decision function being consistent with this. Thus a digital approach represents an attractive method of solving the circuit problems. In principle, the digital type of circuit could be embodied using other active devices, but in practice, owing to the large number of active elements required, transistors have an enormous advantage in power consumption. Furthermore, this is a very suitable first application for transistors, since the demands on their performance are modest, only switching operation being required.

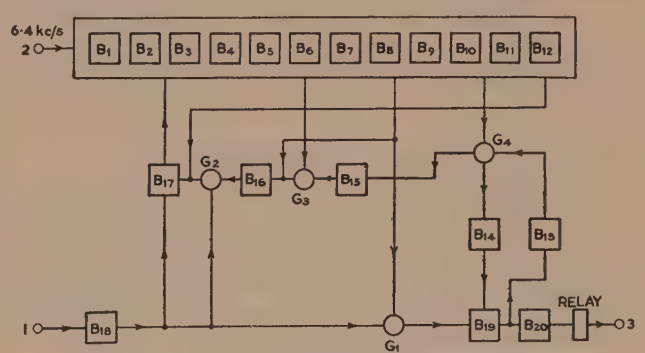


Fig. 1.—Start-stop telegraph regenerative repeater.

Fig. 1 shows a simplified block schematic of a regenerator for 50-baud start-stop transmission. It consists of 20 binary stages of the form shown in Fig. 2 interconnected by gates.

The signal to be regenerated is applied at terminal 1 and B_{18} acts as an input relay. As soon as a mark-to-space transition is received, an impulse is sent by B_{17} to the counting train B_1–B_{12}. The train is normally in a static condition, but when an impulse is received from B_{17}, it starts to count from a 6·4 kc/s signal applied at terminal 2. The 6·4 kc/s signal comes from a crystal oscillator which may be common to a large number of

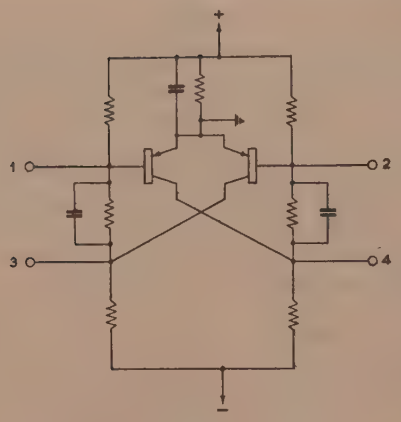


Fig. 2.—Typical binary circuit.

regenerators, and all timing functions are done by successive divisions by two of this signal, together with some extra interconnections to produce timings not given directly by a binary scale. By this means the maximum retiming error occurring in a character can be kept down to less than 0·1 millisec.

Ten milliseconds after starting the counting train an impulse applied to G_1 allows the condition at the output of B_{18} to be transferred to B_{19}. B_{19} holds this condition, which is passed on through B_{20} to the output relay. Impulses are then applied to G_1 at 20 millisec intervals to regenerate the remaining elements in the character. 138·125 millisec after starting the counting train (this figure conforms with C.C.I.T.T. recommendations) an impulse is applied to B_{17}, causing it to stop and reset the counting train. The regenerator is then ready to receive another character.

To carry out the secondary functions the additional binary stages B_{13}–B_{16} and gates G_2–G_4 are used.

(2.1.3) Audio Repeaters.

Although carrier systems have gradually encroached on the field originally exclusively occupied by audio systems, the number of audio links in use is increasing every year. The use of transistors in audio repeaters does not generally permit any very great reduction in first cost or size, since some of the more costly and bulky components are controlled by the need to handle large signalling currents and by protection requirements. However, it is well worth while applying transistors for the following reasons:

(*a*) The power consumption is greatly reduced.
(*b*) The reliability is increased, with consequent reduction of annual charges.
(*c*) It is frequently necessary to limit the gain of audio repeaters to an amount well below the capability of single-stage valve amplifiers. This occurs particularly in the 2-wire and negative-impedance cases, where the usable gain is limited by impedance-balancing problems.[27] In such cases a single-stage transistor amplifier fits the requirement very well.

(2.1.4) Voice-Frequency Telegraphy.

In voice-frequency telegraphy a number of telegraph transmissions may be passed through one speech channel by allocating to each telegraph channel a tone which may be modulated in amplitude, frequency or phase. In Europe, 24 telegraph channels are accommodated in one speech channel, using 120 c/s spacing. Amplitude modulation is in very widespread use;[28] frequency modulation[29] was introduced much later, and for this reason, and also because it has so far been more expensive, is less widely used. However, it has advantages in signal/noise ratio and insensitivity to level variations which make the extra cost worth while in some cases, e.g. on some types of radio circuits. Amplitude and frequency modulation cover the bulk of present systems, other forms of modulation or derived systems being relatively unimportant.

Transistors are well suited to the types of circuit required but fit the f.m. requirement better than the a.m. one, and consequently there is an important change in their relative costs when compared with valve systems. This change arises from two sources, one associated with the transmitting equipment and the other with the receiving equipment.

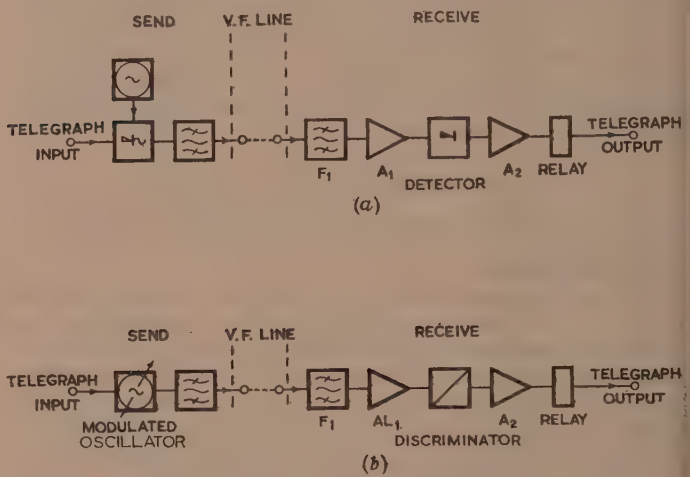


Fig. 3.—Voice-frequency telegraph channels.
(*a*) Amplitude modulated.
(*b*) Frequency modulated.

Figs. 3(*a*) and 3(*b*) show the a.m. and f.m. equipments respectively, an oscillator using one transistor being incorporated in the sending equipment in both cases. In the a.m. case the oscillator output is fed to a modulator controlled by the telegraph signal, and in the f.m. case the telegraph signals are applied directly to the oscillator to control its frequency. Thus, balancing the cost of the modulator needed in the a.m. case against the extra cost of the oscillator needed in the f.m. case, the two equipments cost about the same. The same argument would apply to similar equipment using valves, but while it is essential to have one valve per channel in the f.m. case, a worthwhile economy can be made in the a.m. case by sharing a source of carrier supply between a number of channels. An example of this is the well-established use of motor-generators to supply 10 systems (normally of 18 channels per system). With transistor oscillators, however, the equipment, power and maintenance costs are all a much smaller fraction of the total channel cost. This has the effect of greatly reducing the advantages of shared supplies and therefore narrowing the difference between amplitude and frequency modulation.

The a.m. receiving equipment consists of receiving filter F_1, carrier amplifier A_1, detector, output stage A_2, and relay. It has to select one of the carrier frequencies and to determine the signal instants of the telegraph message, conveyed by the half-amplitude condition. This is determined by the detector, and since only passive elements are needed, similar circuits may be used with valves and transistors. With signals covering a wide level range, the essential problem in the receiver is to have A_1 sufficiently linear at the highest level and still ensure that at the lowest level there is sufficient power to switch the output stage A_2 reliably. With valves it has been found that a single valve in A_1 and another of the same type in A_2 give satisfactory performance. It is necessary, however, to use several transistors in the output stage of A_1 to drive one (or a push-pull pair) of the same type in A_2.

The f.m. receiver consists of receiving filter F_1, limiting amplifier AL_1, discriminator, output stage A_2 and relay. Here the signal instants are determined by the discriminator, which is supplied with a constant level from AL_1 and delivers a constant level to A_2. Thus the performance requirements imposed on AL_1 and A_2 are greatly reduced compared with the a.m. case, and here valves are rather too good and transistors fit the requirements efficiently. In AL_1, transistors will perform the combined functions of limiting and amplifying, since they overload very sharply, whereas with valves auxiliary limiting circuits are needed for a similar performance.

This brief discussion of the essential features of a.m. and f.m. equipment shows that transistors fit more naturally in the f.m. case than in the a.m. case, so that the cost difference between them is reduced compared with valve systems.

(2.1.5) Carrier Telephone Channelling.

In frequency-division-multiplex systems the apparatus in which the first and last modulation processes are carried out (channelling equipment) represents a high proportion of the terminal equipment. The most widely used form of channelling equipment produces the group of 12 channels in the range 60–108 kc/s. Normally the only active circuits particular to each channel are an amplifier after the demodulator in the receiving equipment and, sometimes, a signalling receiver. The main effects of using transistors are as follows:

(*a*) Less heat is produced and more equipment can be installed in a station without forced ventilation.

(*b*) It has been customary to use the power-handling and gain properties of ordinary valves to produce high relative levels (up to +13 dB) at the amplifier output. Such levels are seldom required, but, being within the scope of valves, cause little extra complication and are useful when long audio lines are encountered. The need for high levels may be avoided by having amplifiers at the receiving end of the audio lines, but in the past this has been inconvenient, since it has meant the introduction into exchanges of power supply and maintenance problems not otherwise encountered there. With transistors, significant simplification of the amplifier in the carrier equipment is possible if the relative level at the output is limited to about +6 dB, with additional amplifiers powered by the station batteries at the exchange end of the occasional long lines.

(*c*) Problems of power supply and maintenance technique have also led to a tendency to make signalling receivers part of the carrier equipment rather than of the exchange equipment, even though this may have been inconvenient from other points of view.

For the reasons mentioned in (*c*) the use of transistors permits the location of signalling receivers on more fundamental considerations.

(2.1.6) Three-Channel Open-Wire System.

Three-channel systems providing three carrier telephone channels on an open-wire pair already carrying a 2-wire voice channel were among the earliest carrier systems. Many variants now exist,[30] although these all have the following features:

(*a*) Different frequency ranges are used for the two directions of transmission in order to avoid the intolerable near-end crosstalk which would otherwise occur with open-wire lines. Most systems use frequencies below 19 kc/s for one direction of transmission and above 19 kc/s for the other direction.

(*b*) To provide a reasonable signal/noise ratio under open-wire conditions without making the repeater spacing uneconomically short, the level transmitted to line is high compared with that used on cable systems.

Thus any new systems must use similar frequency ranges and levels to be compatible with existing systems. Unfortunately, while the frequency ranges in use are all broadly covered by (*a*), a great many variants are in use. The existence of these variants combined with customer preferences and special requirements (such as the provision of voice-frequency telegraph channels in addition to three channels of telephony), make it difficult to design a 3-channel system of wide application.

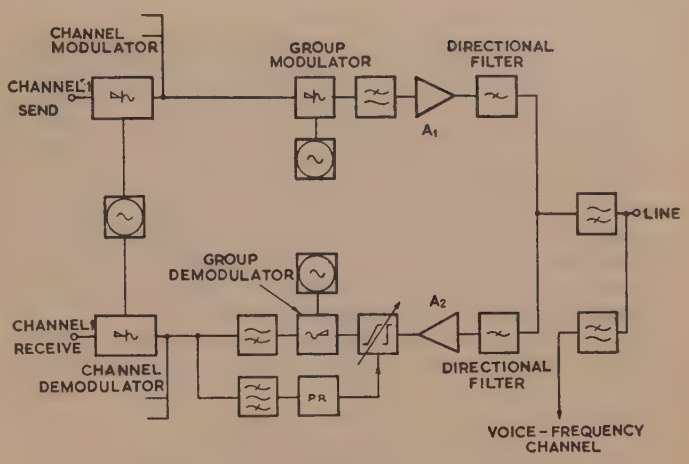


Fig. 4.—Three-channel open-wire terminal.

PR = Pilot regulator.

In the 3-channel terminal shown schematically in Fig. 4 the advantages of transistors are exploited to provide an arrangement in which many variants can be produced by circuit alterations which are simple from the manufacturing aspect. This is achieved in the following way. First, a group of three channels is produced at a frequency well above the line spectrum by using three channels from the standard 60–108 kc/s group mentioned in Section 2.1.5. Secondly, this group of channels is modulated as a group to produce erect or inverted channels at the required line-frequency allocation. This method of producing the line allocation is well known. The advantage of transistors lies in the use of separate crystal-controlled oscillators for each of the carrier frequencies required for the first and second stages of modulation and demodulation. With separate oscillators, different line-frequency allocations can be produced simply by changing the crystals. In principle, of course, valve oscillators could be used, but this would be unattractive, since this is a case where a transistor can supply the small power required efficiently whereas any normal valve would be inefficient. This is an important aspect of 3-channel systems, since they are often used in locations where the main power supply is unreliable and battery operation is very desirable.

In its main features the block schematic is conventional and needs little further explanation. Telegraph channels may be added by injecting them as super-audio telegraphs on any of the telephone channels at the expense of some reduction of speech band in the channels used.

One problem in producing a transistor system which is similar in its external features to existing valve systems is to provide sufficient power-handling capacity in the output stage of the transmitting line amplifier A_1 to transmit three speech channels to line at a relative level of up to +17 dB—the maximum recommended by the C.C.I.T.T. This is achieved by using several transistors in the output stage in a circuit of the type described in Section 6.

Another point requiring more careful attention than in the valve case, though it is not particularly difficult to deal with, is the protection of the transistors in A_1 and A_2 from damage due to transient voltages on the line, caused, for example, by lightning. This problem can be dealt with by adding to the protective precautions normally used on valve systems a limiter consisting of biased diodes or Zener diodes[15] on the equipment side of the line transformer.

(2.1.7) **Twelve-Channel 2-Wire System.**

The 12-channel 2-wire system[31] provides two-way transmission for 12 speech channels over a single pair of wires, thus avoiding the need for separate cables for 'go' and 'return' transmission. This is done by using different frequencies for the 'go' and 'return' directions, the usual allocation being 6–54 kc/s for one direction and 60–108 kc/s for the other. The two bands are separated by means of directional filters at each repeater and are amplified by separate amplifiers. The amplifier outputs are recombined in similar filters for retransmission to line. As already explained, the application of transistors to well-established systems presents some difficulties.

From a study of the block schematic of a transistor repeater shown in Fig. 5 it will be seen that there are two amplifiers in

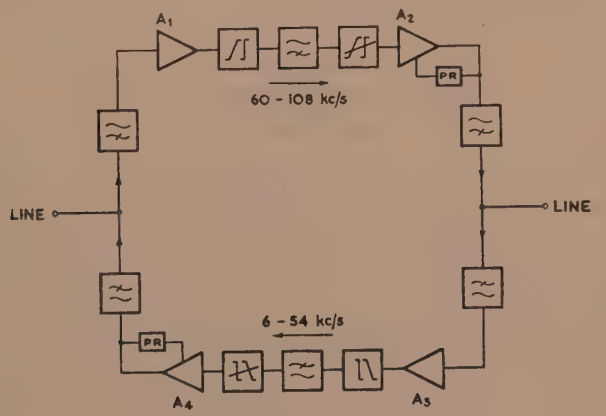


Fig. 5.—Twelve-channel 2-wire repeater.
PR = Pilot regulator.

tandem for each direction of transmission. This feature is necessary in order to achieve with transistors what was previously done with valves. With the smaller signal available from transistors, even greater care must be taken to avoid losses between the output of the amplifier at one repeater and the input to the amplifier at the next repeater. Ideally the only loss here should be that of the cable. With only one amplifier for each direction of transmission (the usual arrangement in valve systems) it follows that any other equipment (filters, equalizers, etc.) must come in the critical position between the output of one amplifier and the input of the next. By dividing each amplifier into an input and an output unit it is possible to put some equipment between the two parts of the amplifier. The equalizer can obviously be moved, but it is also possible to place some of the filtering between the two amplifiers, thus reducing the requirements and therefore the pass-band loss of the main filters. This is because the directional filters have two independent requirements to meet—the efficient separation and combination of the two frequency bands and the prevention of singing round the loop formed by the four amplifiers of the repeater in tandem.

When considering the first requirement, the stop-band attenuations of the individual high-pass and low-pass filters comprising the directional filters need be only moderate. The required frequency band should be attenuated as little as possible, rather than the unwanted band being completely eliminated. It is true that the amplifier loading would be slightly increased by the vestige of the unwanted band, but the extra intermodulation due to this would be negligible. However, if such filters were used, it is certain that the second requirement would not be met and the repeater would sing. In practice, a safe margin is required against singing, so that the individual stop-band losses would need to be considerably increased, the more so because in a long system this effect is additive from repeater to repeater. By dividing the amplifier the extra filtering required can be inserted between the two parts, thus allowing the design of the directional filters proper to be determined solely by the requirement of efficient transmission of the wanted band.

This division of the amplification carries a slight penalty in that it is more expensive to provide a given amount of amplification in two separate amplifiers than in one of the same total gain. About twice as many stages are needed with transistors as with valves, but the penalty becomes small compared with the advantages discussed. An additional advantage of putting some of the filtering between the input and output amplifiers is that some of the filter elements can be reduced in size and cost, owing to the lower level at this part of the circuit. The inter-amplifier circuit is also a convenient point for the application of automatic gain regulation.

(2.2) New Systems using Existing Transmission Media

(2.2.1) **Introduction.**

A further step in the application of transistors is to use them for equipment which would have been impractical or uneconomic using valves. Their most striking advantage is their reduced power consumption, and this makes practical the location at unattended stations or on subscribers' premises of equipment which, in a valve version, would have required too much power for economic battery operation.

A good example of this type of application is the provision of carrier circuits between exchanges and subscribers in rural areas using existing open-wire lines as the transmission medium.[32] From the simplest economic aspect the use of carrier systems for this type of requirement must be compared with the alternative of providing new overhead wires, but, in practice, factors other than straightforward economic ones are important. In particular, carrier equipment held in stock can be rapidly applied when and where the need arises; furthermore, when demand in one area reaches the point where more elaborate systems are justified, the simple systems can be taken out and used elsewhere.

(2.2.2) **Rural Carrier System.**

Apart from the need for low power consumption, systems for use between exchanges and subscribers differ from normal carrier systems in that it is generally necessary to be able to connect individual subscribers to the open-wire line at a number of different points, whereas normally all channels terminate in the same place. Thus, what is needed is really a number of single-channel systems with suitable frequency allocations to fill up the available frequency spectrum.

Fig. 6 shows two frequency allocations based on double-sideband channelling with 8 kc/s spacing of adjacent allocations

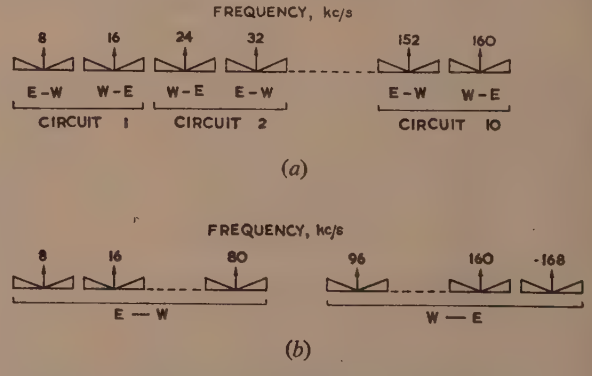


Fig. 6.—Rural carrier system frequency allocations.
(*a*) Stacked.
(*b*) Grouped.

suitable for use with normal open-wire lines. Where there are no problems of compatibility with other open-wire systems on the same route the stacked arrangement would be used. The advantages are threefold. First, since successive pairs of allocations are used for successive channels, the line and equipment can be very efficiently exploited by using circuit 1 to the most distant subscriber, circuit 2 for the next most distant, and so on. Secondly, it allows the system to be built up as the demand increases. Thirdly, sometimes one or more of the frequency allocations may be unusable, because of high attenuation caused by absorption effects or because of noise effect confined to a small part of the spectrum. In these cases the difficulty can be avoided by simply not using the troublesome allocation.

The alternative, 'grouped', allocation, in which one group of consecutive allocations is used for one direction of transmission, is used when crosstalk might arise through conventional systems working over the same pole route.

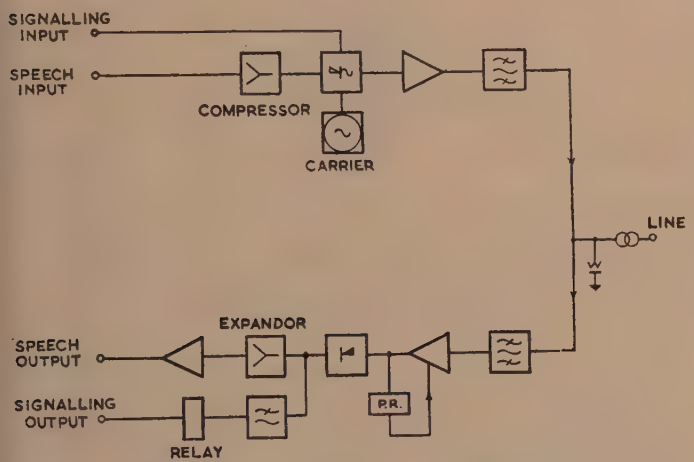


Fig. 7.—Rural carrier out-station equipment.
PR = Pilot regulator.

Fig. 7 shows a block schematic of the equipment needed at the subscribers' end of one circuit for one 2-way transmission. At the exchange end the same equipment would be required, except for the line units, which would be common to all circuits.

On the transmitting side the incoming audio speech is first applied to the syllabic compressor. This reduces the range of levels at its output to roughly half that at its input. The use of a compressor, with the corresponding expandor at the receiving end, improves the signal/noise ratio of the system by at least 20 dB and this permits the use of lines which would otherwise be too noisy and also simplifies the channel filters. The compressor output is applied to a double-sideband transmitted-carrier modulator, the carrier supply for which is derived from an oscillator with coil-and-capacitor frequency control for the lower frequencies and crystal control for the higher frequencies. The modulator output is then amplified and transmitted to line through a band-pass filter and protective equipment (similar to that described in Section 2.1.6), the level of carrier applied to the line being +5 dBm. Signalling is carried out by impulsing the carrier frequency.

On the receiving side the high-frequency equipment consists of a band-pass filter and an amplifier, gain-controlled by the carrier level. The amplifier is designed to deal with a line loss of 28 dB with variations caused by weather from −3 to +6 dB. After rectification the signal is expanded, using a network inverse to that used in the compressor, is amplified and transmitted to the subscriber. The signalling is received by applying rectified carrier through smoothing circuits to a relay.

The transmitting and receiving equipments are constructed as separate physical units, since a variety of sending and receiving frequency allocations may be required in combination. The two units are about the same size, each having a volume of about 200 in^3.

(2.3) New Types of System for New Transmission Media

(2.3.1) Introduction.

The advantages of using transistors already discussed may be further exploited in two important ways. First, line costs can be reduced by using much smaller and more efficient repeaters than would have been possible with valves. A system illustrating this point is described in Section 2.3.2. Secondly, digital equipment becomes practicable, with the possibility of entirely new types of cable using regenerative repeaters. An example of this is the pulse-code-modulation system described in Section 2.3.3. In both these systems two types of transistor are used, having cut-off frequencies of 1 Mc/s (mentioned in Section 2.1.1) and 10 Mc/s respectively.

(2.3.2) 300-Channel Coaxial-Cable System.

Coaxial-cable systems using valve repeaters are well known, systems using frequencies up to 2·6 Mc/s (600 channels)[1] and 4 Mc/s (960 channels)[33] being in general use. These require shorter repeater spacings than earlier carrier systems for smaller numbers of channels (6 miles in the cases mentioned) and therefore the problem of supplying power to repeaters with sufficient reliability is important from the economic standpoint. The usual solution to this problem is to feed the power supply over the transmission cable itself. Another important consideration is that of housing the repeaters, small huts being normally required. The power supply and housing problems are considerably eased when transistor repeaters are used, and it is necessary to reconsider the economics of coaxial-cable systems.[34, 35]

An economic study has shown that a commercially attractive system can be produced using foam-polythene-insulated coaxial cable of 0·163 in diameter with repeaters at 4000 yd intervals, providing 300 channels on a 4-wire basis (C.C.I.T.T. supergroups 1–5). Power is required only at 75-mile intervals—a feature which is necessary in many territories and which would be impractical with valve repeaters.

Fig. 8(*a*) shows a block schematic of the repeater, which is conventional in its general form. Power is supplied by a constant-direct-current feed and the consumption of each 2-way repeater is about 1 watt. The power transmission is separated from the intelligence transmission by the filters marked PSF. The transmission equipment is in two identical parts, one for each direction of transmission. Each amplifier has four transistor stages and has a frequency characteristic compensating the loss of 4000 yd of cable; the gain at the top frequency is 28 dB.

To correct for the variation of cable loss with temperature, a pilot signal is transmitted with the telephony. At each repeater this is picked off from the amplifier output by the selective input filter of the pilot regulator, amplified by three tuned transistor stages and rectified. The rectified output current of the regulator controls the resistance of a thermistor in one of the inter-stage networks of the line amplifier. This network is a variable equalizer, designed to adjust the repeater gain to correct for changes in cable loss with temperature, there being sufficient range to deal with overhead cable. The action of the control loop containing the regulator is to keep the level of the pilot signal at the repeater output constant. The thermistor-controlled inter-stage network is designed so that when the pilot level is kept constant so are all other frequencies in the transmitted band.

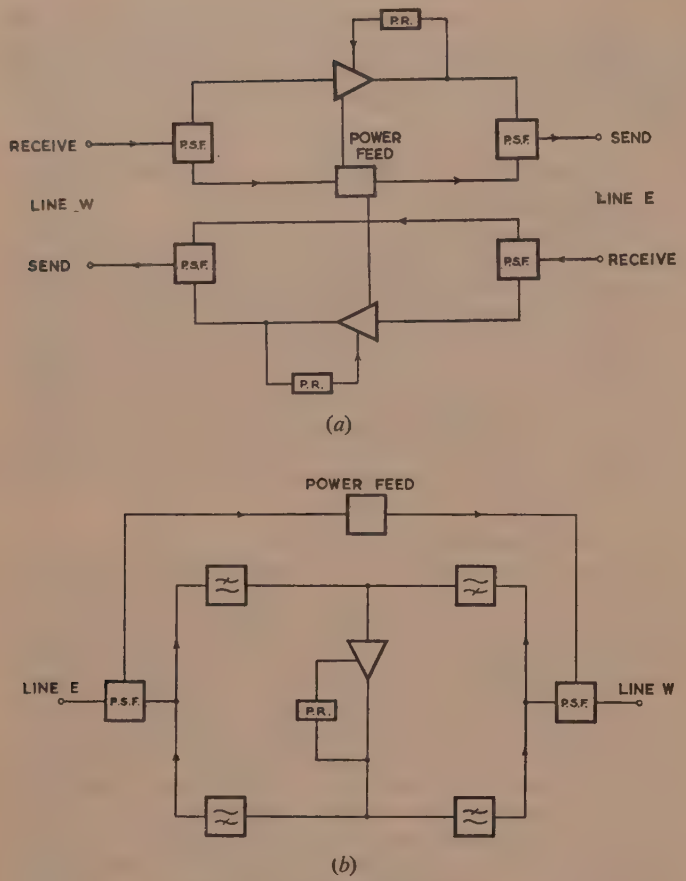


Fig. 8.—Coaxial cable repeaters.

PR = Pilot regulator.
PSF = Power separating filter.
(*a*) 4-wire.
(*b*) 2-wire.

The same basic transmission equipment may be also used on a 2-wire carrier basis to provide 120 channels of 2-way transmission in one coaxial tube. In this case C.C.I.T.T. supergroups 1 and 2 allocation are used for one direction of transmission and supergroups 4 and 5 allocation for the other direction. Fig. 8(*b*) shows the block schematic of the 2-wire repeater, one set of amplifying and regulating equipment handling both directions of transmission simultaneously.

Since the volume of one 2-way repeater is only about 300 in^3 it is clearly unnecessary and uneconomic to provide a building for it. On the other hand, to achieve the long life and high stability needed with so many repeaters in tandem it is necessary to avoid too large or too sudden changes of temperature. As a rough guide to the magnitude of this problem, existing valve systems with repeaters at 6-mile spacing often need special regulating devices to allow for ambient-temperature effects. A solution of this problem is to mount the repeater below ground level in a type of housing permitting rapid access for maintenance. A suitable arrangement is shown in Fig. 9. It consists of a cylindrical metal tube about 9 in in diameter with its bottom end permanently closed and its top end closed by a gasketed cover. The cables enter the housing at the top and are connected internally to the repeater by means of flexible cables. The repeater is normally suspended at the bottom of the housing on a chain, but for maintenance purposes may be drawn to the top and locked in position. With this arrangement the temperature changes are limited to a few degrees, with negligible effect on

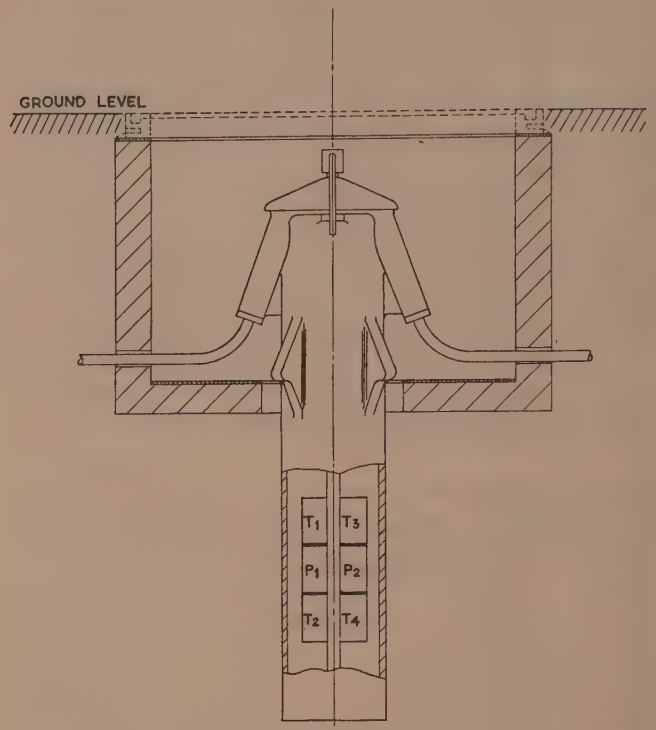


Fig. 9.—Coaxial repeater housing.

transmission, and the best conditions are provided for long transistor life.

In the case illustrated the frame in which the repeaters are suspended is designed to hold two 2-way 4-wire repeaters, and is therefore suitable for a 4-tube cable. Each 2-way repeater consists of three parts, the power equipment, P_1 and P_2, the transmission equipment for one direction, T_1 and T_2, and the transmission equipment for the other direction, T_3 and T_4.

(2.3.3) Pulse-Code-Modulation System.

The idea of transmitting speech by pulse-code modulation (p.c.m.) has been known for many years and experimental systems have been devised.[36–38] The principle is to take a pulse-amplitude-modulated signal and to replace direct transmission of the pulses over the transmission medium by expressing the size of each sample in a telegraph code so that transmission becomes digital instead of analogue. The most common embodiment uses 2-state signals similar to those used in ordinary telegraphy but of much higher speed. Thus, in order to interpret correctly a signal in the presence of noise, it is necessary only to distinguish one of the two states from the other. This makes it possible to use relatively poor cable with simple repeaters.

The main reason why p.c.m. has not been used commercially so far is that the equipment required is more complicated than that for conventional methods of transmission; in particular, it requires a relatively large number of active elements, many of them being used in digital types of circuit. For reasons already discussed, this is an ideal application for transistors and their use totally alters the economic possibilities of p.c.m.

Fig. 10 shows a block schematic of a 24-channel p.c.m. multiplex terminal, those parts which are repeated for each channel being shown only once, the other channels being commoned as shown. With a 10 kc/s sampling rate and a 6-element code, the transmission speed becomes $1{\cdot}44 \times 10^6$ bauds, so that a transistor with a cut-off frequency of the order of 10 Mc/s is required.

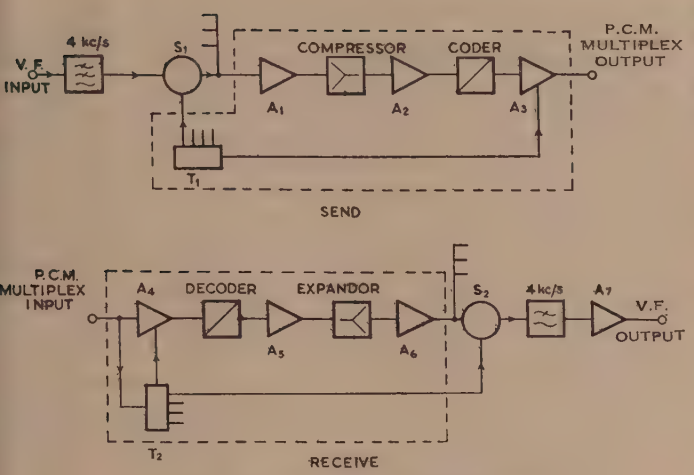


Fig. 10.—Pulse-code-modulation terminal.

S_1, S_2 = Channel gates.
T_1, T_2 = Timing circuits.
A_3, A_4 = Synchronous amplifiers.

The parts enclosed within dashed lines are common to all channels.

The aspects of this equipment which are important in applying transistors are as follows:

(*a*) In the 'send' timing circuit, T_1, it is necessary to generate sampling pulses, equally spaced in time, to open the gates, S_1, in each channel in turn. This is readily done by a counting train consisting of a number of stages in tandem (each having one or a pair of active elements), the rate of counting being determined by a driving oscillator. The 'receive' timing circuit, T_2, can be similar, with the addition of a means of synchronizing the oscillator with the received signal.

(*b*) The amplifiers, A_3 and A_4, can be synchronous regenerative repeaters, their outputs being controlled by signals from the coder in the case of A_3 and from the received line signal in the case of A_4, combined with retiming signals from the appropriate timing circuit. This is a particularly efficient method of operation with transistors, since they dissipate little power internally when switched on or off. This method of use has the additional feature of being relatively insensitive to variations of performance of the active device due, for example, to temperature effects, and therefore makes possible the maximum technical exploitation of a given grade of device.

(*c*) The coder and decoder also use digital types of circuit.

(*d*) The line repeaters are also synchronously regenerative, similar to A_3 and A_4.

Summarizing, there is a strong tendency to require low-power active devices, in many cases operating only as switches, so that small low-power equipment is possible which can be accommodated in housings of the type described in Section 2.3.2.

The robust transmission qualities, small size and low power consumption of such equipment will open up new possibilities in the exploitation of existing cables which are too poor for f.d.m. transmission, and in the introduction of new, less costly, types of cable.

(3) FUTURE

(3.1) Wider Application of Established Types of Transistor

In Section 2 examples are given of the range of work which is possible using two types of transistor, with cut-off frequencies of 1 and 10 Mc/s respectively. In addition to the examples given, transistor versions of most important valve systems up to about 2 Mc/s can be produced using these two types, while new types of system using these transistors can be expected to reduce the minimum distance for which carrier systems become economic compared with audio systems.[39] This should make more practicable the use of carrier systems on exchange junction circuits. An example of a p.c.m. system which may prove economic in junction circuits is given in Section 2.3.3, but there is no doubt that f.d.m. systems will also have applications in this field.

The more obvious attractions of transistor multiplex systems for this type of work (use of exchange batteries, simple maintenance, low-power repeaters) have been discussed elsewhere in the paper but there are a number of other important points, e.g.

(*a*) There is a need to improve the performance of junction circuits. This arises from the high standards set by modern long-distance circuits and from the need to connect an increasing number of links in tandem.

(*b*) Where it becomes necessary to increase the number of circuits between two exchanges beyond the capacity of existing ducts, it is very attractive to apply carrier systems to existing pairs. Alternatively, if there is some space left, more new circuits can be obtained with carrier cable than with audio cable.

(*c*) The freedom in the use of active elements which is possible with transistors is important in dealing with the rather complex signalling conditions which are often met in junction circuits.

(*d*) By combining multiplex systems with electronic exchanges a major step will become possible in the elimination from the overall connection from subscriber to subscriber of the large signalling currents used at present.[40]

Further improvements in reliability combined with the cost reduction brought about by the introduction of new types of component and new manufacturing techniques will accentuate this tendency to use carrier systems for short distances.

(3.2) Future Transistor Development

To extend the scope of transistors beyond the field of application already discussed, improvements are required in frequency response, toleration of high temperatures, power rating, and reliability.

Frequency response is of obvious importance in itself, since the maximum useful frequency is directly dependent on this feature, being about one-fifth of the cut-off frequency in the type of equipment of interest here.

Greater tolerance of high temperatures would permit the extension of transistor applications in three possible ways. First, it would allow transistor equipment to be mounted in places where extremes of atmospheric conditions occur. Secondly, it would make possible further size reduction of equipment without difficulties from internal temperature rise. Thirdly, it would permit a greater power dissipation.

Higher power ratings would be useful both in those cases where high power has to be handled, e.g. to operate electro-mechanical devices, and also in decreasing distortion in multi-channel equipment, assuming that the increased power rating is not obtained at the expense of linearity.

It is unlikely that these improvements can be obtained by further development of the germanium alloyed-junction transistor beyond a cut-off frequency of about 10 Mc/s.[41] The use of an *n–p–n* structure would give a theoretical improvement of about 2 : 1 in frequency response compared with the customary *p–n–p* junction of the same dimensions, but difficulties occur in alloying the *n*-type emitter and collector electrodes.[42] Grown *n–p–n* junctions have a good frequency response and a low collector capacitance, but linearity is poorer, power handling is very limited, and serious troubles occur at high temperatures. The diffused-base transistor,[43] possibly with a high-resistivity region adjacent to the collector to give the effect of a *p–n–i–p*[44] or drift transistor,[45] seems the most likely way of simultaneously extending frequency response and power-handling capacity.

For operation at higher temperatures we may expect some improvements in the use of germanium with advances in processing and encapsulation, but there seems little doubt that silicon will always be superior to germanium in this respect. The lower mobility in silicon is a disadvantage compared with germanium for high-frequency operation, but developments in

diffusion methods have extended the useful range up to about 100 Mc/s.[46] It seems likely that silicon transistors will have many applications in line transmission equipment, although the extent to which they replace germanium transistors will depend on economic considerations.

(3.3) Application of New Types of Transistor

The developments briefly discussed in Section 3.2, combined with steady improvements in other components, may be expected to have the following main influences on equipment:

(*a*) The extension of cut-off frequency and power rating is likely to lead to the elimination of valves from any new designs for land-line transmission within a few years. The most advanced valve system at present in use is the American 8·3 Mc/s coaxial cable system for 1 860 telephone channels.[47] In Europe a 12 Mc/s system for 2 700 telephone channels will shortly be in use. Transistors approaching the performance needed to produce comparable systems are already available in small quantities, and it is to be expected that practical systems will be designed within a few years.

(*b*) The design of present equipment is strongly influenced by the need for access for routine and emergency maintenance. Greater freedom in the use of active circuits will make practicable self-adjusting equipment needing no routine maintenance. Improved reliability will reduce the need for emergency maintenance, and improved temperature tolerance will permit more compact mechanical arrangements. Apart from the direct advantage of reduced maintenance, these factors should lead to simplification and consequent cost reduction. For example, in the case of repeaters, burying in the ground and installation in above-ground cabinets subjected to sun temperatures should become practicable.

(*c*) As discussed in Section 2.3.3, transistors are well suited to digital types of circuit. This fact, coupled with the development of digital technique for computers and the need for transmission between computers, is likely to lead to a greatly increased use of digital transmission, such as pulse-code modulation, for speech.

Summarizing, it is expected that transistors will entirely displace valves in any new designs of land-line transmission equipment within a few years. Great advantages in first cost and running costs will be produced by taking reduced maintenance into account as one of the most important factors in mechanical and electrical design.

(4) ACKNOWLEDGMENTS

The authors are indebted to Standard Telephones and Cables, Ltd., for permission to publish the paper. They also wish to express their gratitude to many of their colleagues who have helped in its preparation.

(5) REFERENCES

(1) Montgomery, A. W.: 'Long-Distance Telephone Communication Systems', *Electrical Communication*, 1948, **25**, p. 386.
(2) Shockley, W.: 'Transistor Physics', *Proceedings I.E.E.*, Paper No. 1941, January, 1956 (**103** B, p. 23).
(3) Loeb, H. W.: 'Transistors versus Thermionic Valves', *A.T.E. Journal*, 1954, **10**, p. 199.
(4) Fink, D.: 'Transistors *v.* Vacuum Tubes', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 479.
(5) Saby, J. S.: 'Fused Impurity *p–n–p* Junction Transistor', *ibid.*, 1952, **40**, p. 1358.
(6) Brockbank, R. A., and Wass, C. A. A.: 'Non-linear Distortion in Transmission Systems', *Journal I.E.E.*, 1945, **92**, Part III, p. 45.
(7) Gartner, W. W.: 'Temperature Dependence of Junction Transistor Parameters', *Proceedings of the Institute of Radio Engineers*, 1957, **45**, p. 662.
(8) Herold, E. W.: 'New Advances in Junction Transistors', *British Journal of Applied Physics*, 1954, **5**, p. 115.
(9) Ebers, J. J.: 'Alloyed Junction Transistor Development', *Bell Laboratories Record*, 1956, **34**, p. 8.
(10) Shea, R. F.: 'Transistor Operation: Stabilization of Operating Points', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 1435.
(11) Guggenbuhl, W., and Schneider, B.: 'Zur Stabilisierung des Gleichstromarbeitspunktes von Flächentransistoren', *Archiv der Elektrischen Übertragung*, 1956, **10**, p. 361.
(12) Tanenbaum, M., *et al.*: 'Silicon *n–p–n* Grown Junction Transistors', *Journal of Applied Physics*, 1955, **26**, p. 686.
(13) Nelson, H.: 'A Silicon *n–p–n* Junction Transistor by the Alloy Process', Transistors I (RCA Laboratories, Princeton, New Jersey, 1956), p. 172.
(14) Shea, R. F.: 'Transistor Audio Amplifiers' (Wiley, New York, 1955), Chapter 2.
(15) Veloric, H. S., and Smith, K. D.: 'Silicon Diffused Junction Avalanche Diodes', *Journal of the Electrochemical Society*, 1957, **104**, p. 222.
(16) Stansel, F. R.: 'Characteristics and Some Applications of Varistors', *Proceedings of the Institute of Radio Engineers*, 1951, **39**, p. 342.
(17) Bargellini, P. M., and Herscher, M. B.: 'Investigations of Noise in Audio Frequency Amplifiers using Junction Transistors', *ibid.*, 1955, **43**, p. 217.
(18) Lo, A. W., *et al.*: 'Transistor Electronics' (Prentice Hall, New York, 1955), Chapter 7.
(19) Dummer, G. W. A.: 'Components for Transistors', *Wireless World*, 1953, **59**, p. 196.
(20) Sharp, L. W. D.: 'Components for Transistor Circuits', *British Communications and Electronics*, 1956, **3**, p. 167.
(21) Weeks, J. R.: 'Metallized Paper Capacitors', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 1015.
(22) McLean, D. A., and Wehe, H. G.: 'Miniature Lacquer Film Capacitors', *ibid.*, 1954, **42**, p. 1799.
(23) Harris, P. O.: 'Plastic-Film Capacitors', *Ericsson Review*, 1954, **31**, p. 56.
(24) Lunchick, A., and Gikow, E.: 'Characteristics of Tantalum Electrolytic Capacitors', *Electrical Manufacturing*, December, 1956, p. 79.
(25) Glover, W. F., and Buckley, S. E.: 'Ferrites Applied to Carrier Frequency Equipment', *Proceedings I.E.E.*, Paper No. 2232 R, October, 1956 (**104** B, Supplement No. 7, p. 541).
(26) Carter, R. O., *et al.*: 'An Electronic Regenerative Repeater for 7½ Unit Start-Stop Telegraph Signals', *Post Office Electrical Engineers' Journal*, 1949, **41**, p. 222.
(27) Llewellyn, F. B.: 'Some Fundamental Properties of Transmission Systems', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 271.
(28) Lloyd, J. A. H., *et al.*: 'A New Voice Frequency Telegraph System' *Electrical Communication*, 1932, **10**, p. 184.
(29) Chittleburgh, W. F. S., *et al.*: 'A Frequency Modulation Voice Frequency Telegraph System', *Post Office Electrical Engineers' Journal*, 1957, **50**, p. 69.
(30) Tonnesen, S. R., and Thornton, A. M.: 'A New 3 Channel Open-Wire Carrier System and its Application in Norway', *Electrical Communication*, 1939, **17**, p. 263.
(31) Reynolds, J., and Jeynes, E.: 'The Design and Field Trial of a 12-Circuit Carrier Telephone System for use on Unloaded Audio Cables', *Post Office Electrical Engineers' Journal*, 1957, **50**, p. 173.
(32) Boyd, R. C., *et al.*: 'Objectives and General Description of the Type P1 Carrier System', *Transactions of the American I.E.E.*, 1956, **75**, Part I, p. 188.
(33) Collier, M. E., and Simpson, W. G.: 'A New 4 Mc/s Coaxial Line Equipment CEL 6A', *Post Office Electrical Engineers' Journal*, 1957, **50**, p. 24.
(34) Hodgson, K. G., and Roche, A. H.: 'Design Factors Influencing the Economical Size and Spacing of Multi-Channel Telephone Repeaters', *Electrical Communication*, 1940, **19**, p. 100.
(35) Weulersse, J.: 'L'avenir du réseau des lignes à grande distance', *Câbles et Transmission*, 1957, **11**, p. 391.
(36) Reeves, A. H.: United States Patent No. 2270070, February, 1942.
(37) Black, H. S., and Edson, J. O.: 'Pulse Code Modulation', *Transactions of the American I.E.E.*, 1947, **66**, p. 895.
(38) Meacham, L. A., and Peterson, E.: 'An Experimental Multichannel Pulse Code Modulation System of Toll Quality', *Bell System Technical Journal*, 1948, **27**, p. 1.
(39) Hoth, D. F.: 'Systems Planning', *Bell Laboratories Record*, 1958, **36**, p. 229.
(40) Deloraine, E. M.: 'Pulse Techniques with Particular Reference to Line and Radio Communication', *Electrical Communication*, 1956, **33**, p. 183.
(41) Statz, H., *et al.*: 'Design Considerations of Junction Transistors at High Frequency', *Proceedings of the Institute of Radio Engineers*, 1954, **42**, p. 1620.
(42) Hunter, L. P.: 'Handbook of Semiconductor Electronics' (McGraw-Hill, New York, 1956), Chapter 4.
(43) Lee, C. A.: 'A High Frequency Diffused Base Germanium Transistor', *Bell System Technical Journal*, 1956, **35**, p. 23.
(44) Early, J. M.: '*p–n–i–p* and *n–p–i–n* Junction Transistor Triodes', *ibid.*, 1954, **33**, p. 517.
(45) Kroemer, H.: 'The Drift Transistor', Transistors I (RCA Laboratories, Princeton, New Jersey, 1956), p. 202.
(46) Tanenbaum, M., and Thomas, D. E.: 'Diffused Emitter and Base Silicon Transistors', *Bell System Technical Journal*, 1956, **35**, p. 1.
(47) Elmendorf, C. H., *et al.*: 'The L3 Coaxial System', *ibid.*, 1953, **32**, p. 781.

(6) APPENDIX: PUSH-PULL OUTPUT STAGE

As pointed out in Section 1, the transistor has a far greater power flexibility than the valve. In amplifier design, for instance, this enables a single type to be used economically in the low-level input stages as well as in the output stages. The range of powers with which one type can deal can be still further extended by using a number of transistors in the output stage, each one providing its share of the output power.

The simplest way of attempting this is to connect the transistors in parallel, but this is not entirely satisfactory because the fairly large spread in parameters leads to unequal sharing of the signal. A push-pull arrangement is always attractive since this gives cancellation of the even-order harmonics. However, this arrangement can lead to unequal sharing of the load in the usual configurations. A circuit which avoids this difficulty is shown in Fig. 11. Four class-A common-base stages are shown, comprising two pairs in push-pull. (The principle is applicable to any even number of transistors.) The common-base stage has a low input impedance and a high output impedance, and the output current is nearly equal to the input current. Thus equal load sharing is achieved by connecting the inputs in series and the outputs in parallel. Push-pull action is obtained by reversing one pair of transistors on the input side and by connecting the output of this pair to the output transformer in the manner shown.

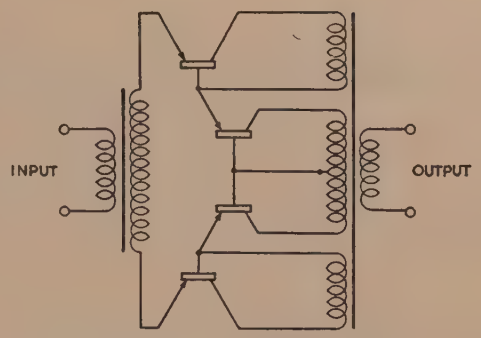


Fig. 11.—Push-pull output stage.

[The discussion on the above paper will be found on page 294.]

621.395.641 : 621.382

The Institution of Electrical Engineers
Paper No. 2764 R
Dec. 1958
©

THE APPLICATION OF JUNCTION TRANSISTORS TO AUDIO-FREQUENCY TELEPHONE LINE AMPLIFIERS

By H. G. BASSETT, B.Sc.(Eng.), and D. THOMSON, Associate Members,
and P. E. GREENAWAY, B.Sc.

(*The paper was first received 3rd May, and in revised form 3rd September, 1958. It was published in December, 1958, and was read before the* RADIO AND TELECOMMUNICATION SECTION *5th January, and the* SOUTH-WEST SCOTLAND SUB-CENTRE *14th April, 1959.*)

SUMMARY

The alloy junction transistor is well suited to application in audio-frequency telephone line amplifiers, where it can supplant the thermionic pentode with advantages in power economy and perhaps in reliability. The application requires good linearity and constancy of gain, and the stability of the input and output impedances is important. The requirements are most simply met by using a common-base circuit without feedback, and a simple amplifier of this type is described; but the use of mixed feedback in a common-emitter stage enables higher gain and output power to be obtained from a given transistor. A single-stage push-pull amplifier using mixed feedback has a performance approaching that of the standard thermionic line amplifier, but it requires individual adjustment of the feedback paths. Nevertheless it has useful applications and is on field trial on a small scale. In order to duplicate the performance of the standard thermionic amplifier without recourse to individual adjustment, two stages are necessary, and two amplifiers of this type are described. The first has a common-emitter stage driving an output stage based on the common-collector configuration, but it requires the use of electrolytic capacitors. A second amplifier, having a common-collector stage driving a common-emitter stage, has only a single paper-dielectric capacitor.

LIST OF SYMBOLS

Z_{in} = Input impedance at the base of a common-emitter stage.
r_e = Emitter resistance in resistive-T equivalent circuit.
r_c = Collector resistance in resistive-T equivalent circuit.
r_{b0} = 'Extrinsic' base resistance.
α = Current gain from emitter to collector.
V_s = Supply voltage.
P_{max} = Maximum available output power.
I_{cb0} = Base-collector current when $I_e = 0$.
β = Current gain from base to collector.

(1) INTRODUCTION

The thermionic amplifiers at present in use in audio-frequency telephone circuits are designed to handle signal levels of the order of 10 mW, and commonly have a gain of 30 or 45 dB. Good linearity and constancy of gain are required, because numbers of amplifiers may be required on long circuits. Input and output impedances are required to be stable and close to 600 ohms, the standard impedance for audio-frequency systems. Present-day thermionic amplifiers generally use mixed feedback, in order to stabilize their gain and output impedance and to reduce distortion to an acceptably low value. Modern amplifiers use indirectly heated valves, the heater consumption of which much reduces the overall power efficiency.

The alloy junction transistor is well suited to application in telephone line amplifiers; most low-frequency units have adequate frequency response, and the medium-power units now becoming available can handle sufficient power to be used in amplifiers which directly replace the thermionic amplifiers.

The authors are in the Post Office Engineering Department.

The relatively high power efficiency of the transistor and its low operating voltage are particularly attractive to the telecommunication engineer; the low operating voltages may well increase component life and can usually be more easily supplied than the high voltages needed for thermionic amplifiers.

Evidence is now available that the alloy junction transistor can be made sufficiently reliable for application in line transmission equipment provided that the operating conditions are suitably chosen and that the transistors are subjected to initial tests to eliminate the few potentially unreliable units.

The work described in the paper concerns the design of some transistor amplifiers suitable for audio-frequency telephony. The amplifiers fall into two classes; those using common-base circuits without feedback, and those using common-emitter, or related, circuits with negative feedback.

(2) A COMMON-BASE AMPLIFIER

(2.1) General

The high linearity of the common-base amplifier, when driven from a constant-current source, enables very simple amplifiers of limited performance to be designed without using feedback. The limited performance arises for two reasons. First, it is usually necessary to provide simultaneously stable, matched impedance conditions at the input terminals and a substantially constant-current drive to the emitter of the transistor. In order to meet these requirements, series resistance must be added in the emitter circuit, with an attendant sacrifice of gain. Secondly, the requirement for stable, matched impedance conditions at the output terminals leads to the provision of a closing resistance across the primary or secondary of the output transformer, with the resulting sacrifice of half the available output power and hence some further loss of gain. Despite these limitations, however, the common-base circuit has been found to have useful applications. Fig. 1 shows a typical circuit including conventional[1] d.c. stabilization.

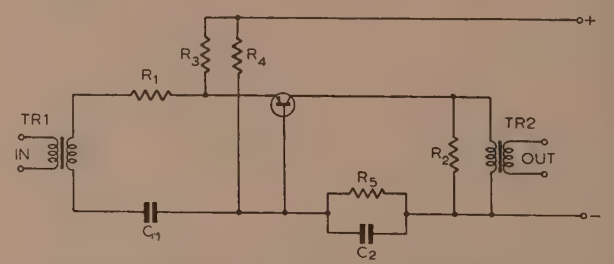


Fig. 1.—An experimental audio-frequency line amplifier using a single common-base stage.

The resistors R_3, R_4 and R_5 stabilize the d.c. operating conditions of the amplifier; R_3 should preferably be of sufficiently high resistance to have a negligible shunting effect on signals at the emitter of the transistor. Too low a value of R_3 gives rise to

excessive distortion owing to the non-linear behaviour of the input impedance, Z_{in}, of the transistor, but, if necessary, this limitation can be avoided by feeding the d.c. emitter bias through the secondary winding of the input transformer (a low value of R_3 can be tolerated). C_1 and C_2 are blocking components, and R_1 is the added series resistance in the emitter circuit. R_2 is the closing resistor for the output.

(2.2) Choice of the Value of R_1

A low value of R_1 is desirable in order to obtain high gain, but too low a value gives inadequate stability of gain and input impedance when the parameters of the transistor change. Increased non-linearity also results from a low value of R_1. Because the collector load impedance is almost invariably small compared with the collector impedance of the transistor, Z_{in} can be assumed to be approximately equal to $r_e + (1 - \alpha)r_{b0}$, and variations in the input impedance of the amplifier can be approximately calculated by investigating the changes in Z_{in} when α and r_{b0} vary over the spreads permitted by the transistor specification. R_1 can then be chosen so as to meet the specified limits upon the return loss of the input impedance and upon gain. Distortion produced by the non-linearity of r_e can be calculated by using the analyses given by Jones[2] and Meyer,[3] but a detailed investigation has not been made because the values of R_1 necessary to give an adequately stable input impedance are found to give more than adequate linearity.

(2.3) The Choice of the Load Impedance

In the present application the load impedance is always small compared with r_c, and is dictated by the available supply voltage and the required overload point of the amplifier. Thus if the collector supply voltage is V_s and the maximum available output power to line is to be P_{max}, the collector dissipation must, in principle, be $4P_{max}$, and hence the load impedance (equal to $R_2/2$) must be $V_s^2/4P_{max}$.

In practice P_{max} must be increased above the design value sufficiently to enable the desired output power to be obtained despite transformer losses and possible variations of bias conditions.

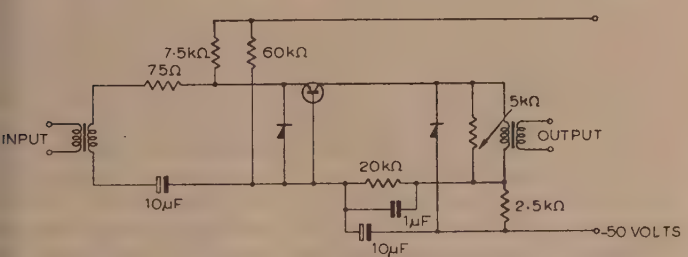


Fig. 2.—A line amplifier using the common-base circuit.

(2.4) Details of a Practical Amplifier

Fig. 2 shows the circuit diagram of an amplifier which has been used in small numbers in the telephone network.

The main design parameters are as follows:

Mean collector-base voltage	10 volts
Mean emitter current	4 mA
Nominal overload point	+10 dBm
Power consumption	50 volts at 4·1 mA
Gain	10 dB
Nominal input and output impedances	600 ohms
Temperature coefficient of gain	<0·002 dB per 1° C
Maximum possible change of overload point due to 100 μA change of I_{cb0}	1·2 dB

The two diodes shown in Fig. 2 are provided in order to protect the amplifier from impulses due to signalling conditions. The diode connected between the emitter and base prevents the emitter from being driven unduly negative relative to the base, and the diode connected to the collector of the transistor catches the collector potential at twice the mean value. Any energy fed into the amplifier by signalling impulses is thus prevented from dissipating itself in the transistor. A typical figure for harmonic distortion in the amplifier is 42 dB (signal/harmonic ratio) for total harmonics at an output level of +6 dBm; the linearity is therefore adequate for most audio-frequency line applications.

(3) THE APPLICATION OF FEEDBACK TO COMMON-EMITTER AMPLIFIERS

(3.1) General

The common-emitter circuit is well suited to applications involving negative feedback,[4,5,6] because it provides a phase reversal between base and collector and also, if the load impedance is suitably chosen, large gains of current and voltage simultaneously. These features are necessary if a useful amount of negative feedback is to be applied without using transformers in the feedback path.

(3.2) Mixed Feedback

The need for stabilized input and output impedances in telephone line amplifiers makes it desirable to employ mixed feedback; it is then possible to stabilize the output impedance of the amplifier against changes of transistor parameters without recourse to the dummy load used in the amplifier described in Section 2. Mixed feedback may be readily applied to single common-emitter stages in two ways [Figs. 3(*a*) and 3(*b*)].

In Fig. 3(*a*), R_s feeds back to the input circuit a voltage proportional to the emitter current (and hence substantially proportional

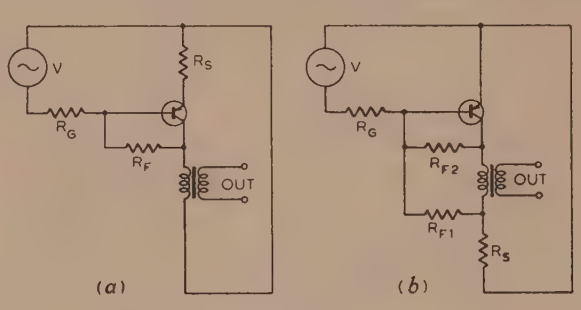


Fig. 3.—Two methods for the application of mixed feedback to a common-emitter stage.

to the collector current), and R_F feeds back a current proportional to the voltage across the load impedance. With a suitable value of R_G, mixed feedback occurs and the output impedance can be readily shown to tend towards $R_s(R_F + R_G)/(R_s + R_G)$ as the feedback is increased. In the alternative circuit, satisfactory operation requires that the signal levels at the right-hand ends of R_{F1} and R_{F2} be large compared with the signal level at the base of the transistor, and that R_{F1} and R_{F2} be sufficiently large for their shunting effects on the collector circuit to be small. These conditions can be approximated sufficiently well in single-stage amplifiers to enable useful feedback amplifiers to be designed. It is readily shown that the output impedance tends towards the value $R_s(R_{F1} + R_{F2})/(R_s + R_{F1})$ as the feedback is increased. It is seen that the expression is independent of R_G. Hence the output impedance of the amplifier can be made almost independent of conditions at the input of the stage, making the

circuit preferable to that of Fig. 3(*a*) for single-stage amplifiers. The impedance R_G can be divided into two parts, one of which is the true generator impedance, to provide matched conditions (see Fig. 4).

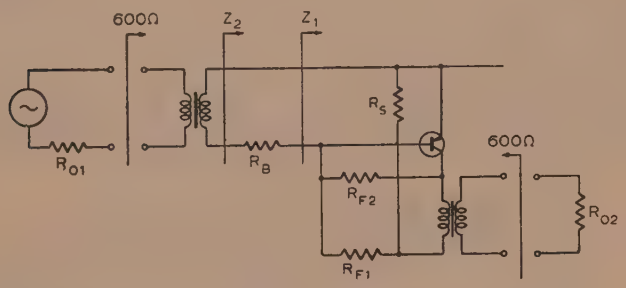


Fig. 4.—The circuit of Fig. 3(*b*) shown in more detail.

As the feedback is increased, Z_1 tends to decrease. In practice, therefore, R_B is chosen to give the required constancy of input impedance, Z_2, against changes of transistor parameters. Too large a value of R_B results in undue sacrifice of gain. Finally, a transformer is provided at the input of the amplifier to make the actual impedance at the terminals equal to 600 ohms; R_{01} can then be 600 ohms also. At the output terminals of the amplifier a transformer is also provided in order that R_{02} can be 600 ohms.

(4) A SINGLE-STAGE PUSH-PULL LINE AMPLIFIER BASED ON THE CIRCUIT OF FIG. 4

Fig. 5 shows a push-pull single-stage line amplifier based on the circuit of Fig. 4.

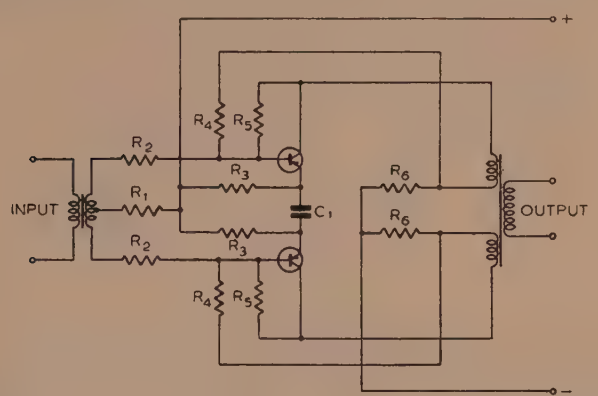


Fig. 5.—A single-stage push-pull line amplifier based on the circuit of Fig. 4.

The circuit makes use of the feedback resistors to form part of the d.c. stabilizing system. Thus the base potentials are prescribed by the potentiometer network R_1, R_2, R_4, R_5 and R_6, and the emitter current is defined by R_3. The resistances R_5 provide shunt feedback and R_4 and R_6 provide series feedback. R_2, corresponding to R_B in Fig. 5, defines the input impedance of the amplifier when viewed from the secondary winding of the transformer. The amplifier needs only a single decoupling capacitance, C_1. Numerical values for the components have not been given in Fig. 6 because the amplifier has been used in two versions, one operating from 24-volt and the other from 50-volt supplies. The main parameters of the two types are as follows:

	24-volt version	50-volt version
Gain	24 dB	27 dB
3 dB bandwidth	150 c/s to >10 kc/s	150 c/s to >10 kc/s
Nominal overload point	+16 dBm	+16 dBm
Mean collector-base voltage	12 volts	25 volts
Mean emitter current	3·3 mA	1·6 mA
Total power consumption	210 mW	240 mW
Nominal input and output impedances	600 ohms	600 ohms
Total harmonics at an output level of +14 dBm	4%	4%
Nominal loop gain when $\beta = 40$	8 dB	4 dB
Temperature coefficient of gain	≯0·025 dB/deg C	≯0·025 dB/deg C
Maximum change of overload point possible due to 100 μA change of I_{cb0}	0·5 dB	0·6 dB

Individual adjustment of the feedback resistances was necessary in order to secure correct output impedance and gain. It was made by adjusting R_5 and R_6 (Fig. 5); R_5 was sufficiently large and R_6 sufficiently small for the required range of variation to result in little change of d.c. conditions.

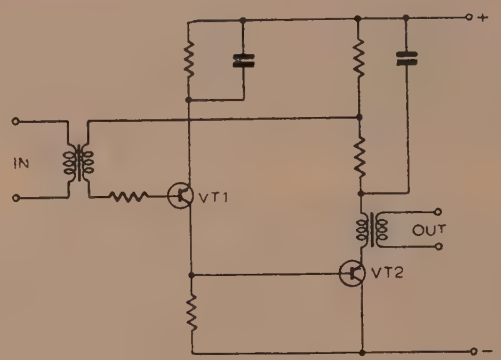


Fig. 6.—Basic circuit of two-stage amplifier showing d.c. stabilization.

Using a given type of transistor, the power output obtainable from the push-pull amplifier is about four times that given by the amplifier described in Section 2, and gains of up to about 30 dB can be obtained. However, it was not possible, using available transistors, to obtain a gain of 30 dB and at the same time apply enough feedback to obviate the need for individual adjustment of the amplifiers. In order to duplicate the performance of one commonly-used standard valve amplifier, therefore, at least two stages are necessary. The designer must decide between a push-pull output stage and a single-sided output stage. Although a push-pull stage gives less distortion than a single-sided output stage, the feedback necessary for adequate stabilization of output impedance is sufficient to give an acceptable performance in respect of harmonic distortion using a single-sided stage. The latter was therefore preferred on grounds of simplicity, particularly as a suitable transistor was becoming available.

(5) A TWO-STAGE LINE AMPLIFIER WITH MIXED FEEDBACK

If feedback is to be applied over two stages, only one of them must produce a phase reversal unless a transformer is included in the feedback path. A practicable circuit arrangement consists of a combination of a common-collector and a common-emitter stage. The amplifier described in this Section is based on a common-emitter stage driving a common-collector stage; an amplifier based on the alternative arrangement is described in Section 6. Considerable economy in components may be

effected, and improved performance obtained, by directly coupling the two stages and using d.c. feedback to stabilize the d.c. operating conditions. Fig. 6 shows the basic circuit before a.c. feedback is applied.

The circuit shown in Fig. 6 suffers from the defect that the output transistor cannot be brought close to the bottomed condition when VT_1 is cut off, because the base current of VT_2 flowing through the load impedance of VT_1 (which cannot, for other reasons, be less than a few thousand ohms) causes a voltage drop which can amount to several volts. An inordinately large dissipation in VT_2 would then be necessary for a given output power. The disadvantage can be overcome by splitting the primary winding of the output transformer and including part of it in the collector circuit of VT_2, as shown in Fig. 7.

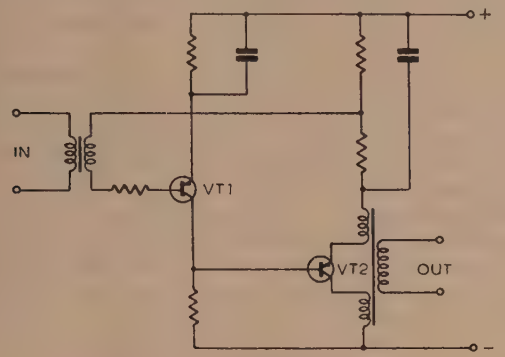


Fig. 7.—Amplifier with modified output stage allowing VT_2 to be fully driven.

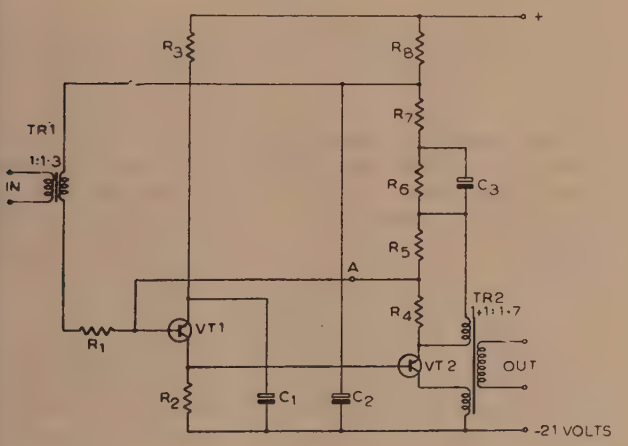


Fig. 8.—Complete diagram of two-stage amplifier type TB4.

Mixed feedback is applied from the emitter circuit of VT_2 to he base of VT_1, leading to the final arrangement shown in Fig. 8.

The a.c. feedback is applied by means of R_7, R_4 and R_5; R_6 nd R_8 are by-passed at signal frequencies, but they form part f the d.c. stabilizing circuit. If we consider the feedback path o be disconnected at A in Fig. 8 for the present, the potential t the right-hand side of the break consists of a portion of the oltage across the primary of the transformer (depending on R_4 nd R_5) plus the voltage across R_7, which is proportional to the utput current. When the feedback path is reconnected at A, herefore, a current (containing components proportional to output voltage and to output current) is fed back to the base of VT_1 and mixed feedback is thus applied. The important parameters of the amplifier are as follows:

Gain	30 dB
3 dB bandwidth	100 c/s–25 kc/s
Nominal overload point	85 mW (+19 dBm)
Mean collector-base voltage of VT_2	12 volts
Mean collector current of VT_2	14 mA
Mean collector-base voltage of VT_1	7·4 volts
Mean collector current of VT_1	1·3 mA
Total power consumption	320 mW
Signal/total harmonics at an output level of +17 dBm	46 dB
Nominal feedback factor when $\beta = 40$ for both transistors	12 dB
Return loss of imput and output impedances against 600 ohms	>15 dB over range 300–3 000 c/s >20 dB at 800 c/s

(6) A TWO-STAGE LINE AMPLIFIER WITHOUT ELECTROLYTIC CAPACITORS

The circuit of Fig. 8 includes three electrolytic capacitors, the reliability and life expectancy of which have yet to be proved high enough for line amplifier installations. An alternative design which avoids the use of such capacitors has been developed.

The function of the capacitors in Fig. 8 is to decouple the resistors which define and stabilize the d.c. operating conditions of the transistors against changes in temperature and spreads in manufacturing tolerances of transistor parameters.

In an unstabilized amplifier in which the operating conditions of the output stage have been chosen (at normal temperature) to give maximum undistorted power into the required load, the effect of a temperature increase is twofold. First the mean collector current increases, causing a change in the operating conditions, and the overload point is lowered; secondly the collector dissipation increases and the temperature of the collector junction rises. The main problem in the design of an amplifier, using a given transistor in its output stage and without using decoupling capacitors, is to provide a stabilizing circuit which will ensure that, over the expected range of ambient temperature (0–45° C for this application) and taking into account the specified spreads in transistor parameters,

(*a*) The collector dissipation in the output transistor will not exceed that permitted.
(*b*) The stabilizing circuit will not consume so much signal power that sufficient output power cannot be delivered to the load.
(*c*) The operating conditions of the output transistor do not alter so much with temperature, or change of transistor, as to reduce the overload point below the specified value.

The d.c. circuit of the amplifier is illustrated in Fig. 9.

The circuit has been developed using an output transistor VT_2 which has a maximum permissible collector dissipation of 250 mW and a maximum peak voltage, collector–emitter, of 30 volts. The input transistor VT_1 has a maximum collector

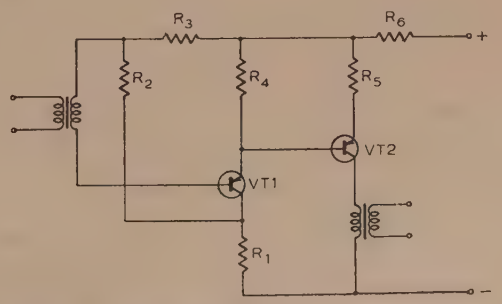


Fig. 9.—D.C. stabilizing circuit of TC1 amplifier.

dissipation of 25 mW and a maximum peak voltage of 30 volts. In Fig. 9 the d.c. operating conditions of VT_1 are fixed and stabilized by the chain of resistors R_1, R_2 and R_3 together with R_4. The base of VT_2 is directly coupled to the emitter of VT_1 and R_5 defines the current in VT_2.

For adequate stabilization, the normal direct voltage across R_4 must be much larger than the base-emitter voltage of VT_2. On the other hand R_5 must not be too large if it is to remain un-bypassed, otherwise the gain of the output stage will be reduced by local feedback. Because R_5 and R_6 are part of the load on the output transistor they must be small compared with the total load impedance.

The operating conditions of the output stage were therefore chosen to utilize the maximum permissible voltage between collector and emitter so as to give the highest possible load impedance for maximum undistorted output power. The ratio of the output transformer follows. By connecting the junction of R_2 and R_3 to the end of the input transformer remote from the base of VT_1, the input circuit of the common-collector stage suffers no shunt loss. In the final circuit the required image-impedance conditions at input and output are obtained by applying negative feedback over the two stages, and choosing a suitable turns ratio for the input transformer. With all the negative feedback removed, the amplifier gain and input and output impedances vary widely with spreads in transistor parameters, temperature and variations of supply voltage. The mean gain at 20° C is about 53 dB. The negative feedback applied is sufficient to stabilize the gain and impedances to the limits given below, without undue restrictions on permissible transistor parameter spreads. For example, spreads of 40–100 at 20° C for the common-emitter current gains of both transistors are permissible. The practical circuit is shown in Fig. 10.

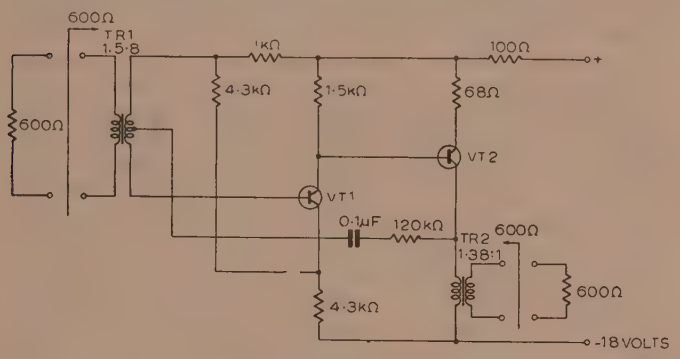


Fig. 10.—Complete diagram of two-stage amplifier type TC1.

The main features of the amplifier performance are as follows:

Operating ambient temperatures	0° C to +45° C
Insertion gain at 1 kc/s	30 dB ± 0·5 dB
Insertion gain at 200 c/s and 8 kc/s.. ..	Equal to gain at 1 kc/s ±0·25 dB
Input and output impedances expressed as return loss against 600 ohms	At 1 kc/s, 20 dB At 300 c/s and 4 000 c/s, 15 dB
Output distortion	5% at 50 mW output
Power consumption (with average transistors at 20° C)	16 mA at 18 volts 290 mW (approximately)
Average collector dissipation of output stage	200 mW
Maximum collector dissipation	250 mW

(7) CONCLUSION

Junction transistors can supplant the thermionic valve in most audio-frequency applications in line transmission. The requirements for gain/frequency response, and for stability of gain and of input and output impedances, can be met by using low-frequency alloy transistors in fairly straightforward circuits employing mixed negative feedback. Mass-produced transistors can be used, and no selection is necessary provided that they meet a specification prescribing tolerances comparable with those of the thermionic valve hitherto used in transmission equipment.

Field trials suggest that high reliability should be obtained at moderate ratings; over the past two years $1{\cdot}9 \times 10^6$ transistor hours have been run, with only two failures, on amplifiers of the type described in Section 4. The mean collector temperatures during the tests were mostly less than 45° C.

(8) ACKNOWLEDGMENT

Acknowledgment is made to the Engineer-in-Chief of the Post Office for permission to make use of the information contained in the paper.

(9) REFERENCES

(1) SHEA, R. F.: 'Principles of Transistor Circuits' (Wiley, 1953), p. 102.
(2) JONES, D. H.: 'Intermodulation Distortion in Experimental Group Amplifiers using Transistors' (to be published).
(3) MEYER, N. I.: 'Non-Linear Distortion in Transistor Amplifiers at Low Signal Levels and Low Frequencies', *Proceedings I.E.E.*, Monograph No. 209 R, November, 1956 (**104** C, p. 208).
(4) ALMOND, J. A., and BOOTHROYD, A. R.: 'Broadband Transistor Feedback Amplifiers', *ibid.*, Paper No. 1929 R, January, 1956 (**103** B, p. 93).
(5) HANNINGSBERG, C.: 'Amplificateur a transistrons à faible consommation', *Onde Électrique*, 1955, **35**, p. 264.
(6) MOULON, J.: 'Propriétés essentielles des transistrons', *ibid.*, 1955, **35**, p. 243.

DISCUSSION ON THE ABOVE TWO PAPERS BEFORE THE RADIO AND TELECOMMUNICATION SECTION, 5TH JANUARY, 1959

Mr. R. H. Franklin: In Paper No. 2722 the authors list the relative advantages and disadvantages of transistors compared with valves. But we want to use transistor amplifiers in repeater stations for very simple reasons. In small stations they enable us to avoid the use of expensive power plant. In large stations we can get more equipment into our repeater-station buildings, some of which inevitably are too small. The other advantages given in the paper may be important, but they are quite secondary for current applications.

There are, of course, great advantages to be gained by using transistors for power-fed line repeaters, because of the reduced power consumption. D.C. instead of a.c. power-feeding eliminates many power units, and the line voltage can be kept low and non-lethal. It is hoped that it will be possible for jointers to work on cables while live and that it will be unnecessary to shut down all the systems in a cable while faults are being localized and cleared.

The paper has one or two controversial statements, as one would expect and hope. In Section 3.1, it is stated that there is need to improve the performance of junction circuits. Presumably this is to be done by carrier working; but surely a pair of copper wires with the loading coils in the right place will give better and more reliable transmission over short distances than we are likely to obtain with any carrier system in the foreseeable future.

In Section 2.3.3, it is suggested that pulse-code modulation

ill give 24 circuits on cables too poor for frequency-division ıultiplex, but I cannot think of any cable now laid that would be oor for 24-channel frequency-division multiplex. Pulse-code-ıodulation will have to prove itself on cost, and it will do this ith no difficulty for many applications.

In Paper No. 2764, the TC1 amplifier unit described is now ıe Post Office standard and will be used in all a.f. 2-wire and -wire repeaters purchased from now onwards. I regret that no ıention is made of the negative-impedance repeater, which is ure to play an important part in providing a.f. circuits until the ulse-code-modulation system is introduced.

Mr. J. R. Cannon: Although I am in sympathy with the view hat the number of types of transistors used in designing a range f equipment should be restricted, if possible to one type, the mportance to the customer can be exaggerated. If transistors re as reliable as we hope—and figures of the order of 100 years ave been mentioned—replacement will be required very ıfrequently.

I am glad that the authors recommend an output level of bout +6 or +7 dB at the 4-wire point. This is a very desirable tep in the direction of designing around the natural properties f transistors instead of merely replacing valves by transistors.

There are difficulties of compatibility, but there is little doubt hat, as experience is gained, new electrical and mechanical tandards will have to be established to make use of the peculiar roperties of transistors. Present mechanical designs seem to e based principally on the necessity to provide an accessible latform on which to mount valves. If we do not have to ıount valves, a new approach is possible.

In Paper No. 2764, the authors have gone to some trouble to void using a transformer as a phase reverser in the feedback ath of their 2-stage amplifier. We have found no difficulty in his approach, and have produced a successful demodulator mplifier, using common emitter in both stages with d.c. coupling etween the two stages and with a.c. and d.c. feedback overall. his configuration permits the use of very small transformers and ives a performance similar to that of the last amplifier men-ioned in the paper.

Mr. S. A. Phillips: Except where they offer special advantages, ransistor systems will need to be competitive in price with orresponding valve systems. On the average there are 2–2½ ransistor stages to each valve stage. Hence economies must e made elsewhere, and certainly the cost of the transistors hemselves will need to be reduced to the order of a shilling or wo. There are other compensating factors, such as a much ower power drain and reliability, but many customers are not nfluenced by running costs.

The first of the rural carrier systems was the Bell P1. This is n expensive system, which, nevertheless, has many imitators in he United States, because of the operational facilities it affords. hese can be realized to the full only by the use of transistors.

The rural carrier system described in Paper No. 2722 is much nore ambitious. Field experience with a system almost identical ith that described by the authors, but manufactured by another ritish company, may be of interest. This equipment has now een in service in various parts of the world (ranging from the ropics to this country) for over 18 months, and has been almost ault free and without any transistor faults whatever. Transistor requency-shift telegraph systems have been running for some our to five years, also without a single faulty transistor.

The 300-circuit system described uses a regulator at every epeater, i.e. every 4 000 yd, but this does not seem justified for uried cables. Perhaps the authors would indicate the advan-ages they gain from this arrangement.

Paper No. 2764 describes an a.f. amplifier without any ecoupling capacitors. Will these amplifiers work from typical repeater-station battery supplies, or are decoupling capacitors required on the power supply to each amplifier?

In conclusion it has been said in jest, but with a modicum of truth, that if the transistor had been invented first, the invention of the valve would have been hailed as the solution to the inherent limitations of the transistor.

Dr. J. R. Tillman: Paper No. 2722 claims improved life for the transistor. But Waltz's* data for the failure rates of seven types of transmission valve (four types had removal rates of less than 0·05% per 1 000 hours and the poorest had a rate of only 0·3% per 1 000 hours) and of some transistors (grown units in the P-type carrier system showed a rate of 0·67% per 1 000 hours and alloy units in a line concentrator 0·2% per 1 000 hours) hardly bears out the claim, particularly if two transistors are normally required to replace one valve. No opportunity of collecting field data on this vital point should be missed.

The claim for greater linearity warrants detailed investigation in terms of circuit configuration and ease of applying feedback.

I do not subscribe to great economy of types if the numbers used are considerable. A factor of three or four, for power-handling capacity or bandwidth, usually justifies a new type.

What parameters of transistors, besides speed of response, limit their use in p.c.m. systems? What are the potentialities of transistors as modulators? Will transistors find uses in reducing the loss and attenuation distortion of filters?

In Paper No. 2764 mention should have been made of modern electrolytic capacitors. Is there not evidence that capacitors made from very pure aluminium, processed to be free of corrosive substances and with properly made terminations and seals, have lives of many years under the favourable conditions of use in the amplifier described in Section 5, namely low steady direct voltages, lack of ripple and non-association with hot components?

The results for the amplifiers mentioned in Sections 4, 5 and 6 are not presented uniformly, which makes comparisons difficult. To what extent are any differences a function of the components, e.g. transformers, other than the transistors? Does one design permit more deterioration of the transistor parameters (particularly I_{CB0} and h_{fe}) than the others?

Dr. J. A. Pim: The British Post Office are fortunate in that their designs of amplifiers need only cater for a temperature range of 0–45° C. For those who intend to design transistorized equipment for export to tropical countries, it is necessary to consider temperature ranges up to +55° C. The difficulties then seem to increase exponentially rather than linearly because of the low maximum junction temperature of most germanium transistors.

It is necessary in many instances to reduce the power dissipation of the individual transistor by increasing the number in the circuit and thus to provide at the same time the extra degree of gain and stabilization required to offset deterioration in performance of the transistors working near their maximum rated junction temperature.

The main stress in Paper No. 2722 has been on the replacement of valves by transistors in line-communication equipment mainly built to C.C.I.T.T. recommendations. These recommendations have, in general, grown up around the capabilities of the normal range of receiving valves, such as the high-gain pentode type CV138. Have we now reached the stage where the levels used in line-transmission systems can, with advantage, be fundamentally replanned round the optimum capabilities of transistors, now accepted as reliable devices, even though such new systems, although economic using transistors, might be uneconomic compared with existing systems if both were using valves?

The use of transistors as circuit-elements in the design of filters and modulation has not been mentioned in the papers. Definite

* 'Transistor Technology' (Van Nostrand, 1958), Vol. 3, chapter 12.

technical advantages can be shown to exist when transistors are used in place of semiconductor diodes in modulators. For a standard level of input signal, the carrier-frequency power required is some 10 dB less than that required for existing diode modulators. This saving in carrier power is obtained without the application of external d.c. power to the modulator.

There is also considerable improvement in fundamental carrier leak. This simplifies the design of sideband-selection filters and facilitates the introduction of out-of-band signalling frequencies in the immediate vicinity of the carrier frequency.

Mr. P. P. Eckersley: The designer must be as much concerned with the cost and reliability of the large capacitors, which usually have to be used in the associated circuits, as with the transistors themselves. Tacit agreement with this statement is contained in Paper No. 2764, which describes an amplifier using only one capacitor. Does it really work without further decoupling, and, more specifically, what would be the limiting value of the source impedance? We seem to be between the devil of high cost and the deep sea of unreliability. A few years ago it was forecast that the development of the small aluminium electrolytic capacitor with low working voltage would meet the designers' needs if attention were directed towards purity of material. In fact, this has not yet been achieved, and, in spite of increased cost, the probable life of such capacitors is too limited before there is an appreciable risk of their becoming open-circuited.

If C is the capacitance in microfarads and V the working voltage in volts, an acceptable specification would be a CV value of 300 at a cost of about 6s. with a life of 20 years. Is such a specification likely to be met in the near future either by aluminium or tantalum types? However reliable transistors may become, their uses will be severely restricted unless the reliability of the capacitors essential to their functioning is greatly improved.

Paper No. 2764 refers to $1{\cdot}9 \times 10^6$ transistor hours as a criterion of reliability. What was the number of transistors used in tests to produce this remarkably encouraging figure?

Mr. E. R. Thomas: In Paper No. 2722 it is suggested that the types of transistors should be limited. This economy could presumably be rather restricting in some cases, particularly in regard to power output if a transistor of only, say, 50–100 mW dissipation is allowed.

The maximum levels quoted for the 3-channel amplifier are maximum test levels, and although no C.C.I.T.T. recommendation is given, some allowance should surely be made for overload. For a 12-channel amplifier the C.C.I.T.T. recommendation is +33 dBm, and something comparable would presumably be needed for the 3-channel amplifier, but a 50 mW transistor would require at least 40 transistors in the output stage. Are the authors not obtaining an overload point of this order, and, if they are, how is this achieved?

In Section 1 it is stated that less feedback need be applied to a transistor amplifier than to a valve amplifier. This must depend on the particular conditions, and the reference here is presumably to the common-base configuration.

In Section 1.2 it is stated that there is little to choose between the common-emitter and common-base multi-stage amplifiers, so far as the overall performance is concerned. The common-base amplifier requires interstage transformers if any gain is to be obtained. Transformers have the disadvantage of high cost, and they should be reduced to a minimum. In addition, they are difficult components to include in a feedback loop because the characteristics can produce added stability problems.

From Table 1, the saving in power consumption using transistors is impressive, namely 10 : 1, but I imagine that this could be still further improved. One difficulty seems to be that of having to operate transistor terminal equipment from standard 24-volt and even 50-volt batteries. One certainly does not need to have a 50-volt battery so far as the transistor is concerned. This can mean that, even after allowing for decoupling, there is still surplus voltage to lose, presumably in dissipating resistors. This seems to be brought out in Paper No. 2764, in which the authors give a series of amplifier designs together with consumption and supply voltage. There is a trend for the overall efficiency to increase as the operating voltage is reduced from 50 to 18 volts.

Mr. G. King: The following data on transistor reliability were put forward at Washington in October, 1958.

Dr. Ebers showed a slide plotting life at 85° C junction temperature against year of manufacture, and, in 1950, lives of a few hundred hours were shown for plastic encapsulated types. Hermetic sealing showed a rise, steeply after 1954, to a figure of some 20 000 hours in 1956. The curve then had a dotted extrapolation to 200 000 hours in 1960, making use of further detailed advances in technique.

Another slide was shown plotting the failure rate against time, for n–p–n alloyed transistors. In this curve, the peak value was 1·5% at about 400 hours, falling to 0·3 and 0·03% (all per thousand hours) at 2 000 and 6 000 hours. A further slide showed the power ageing of some of the newer transistors run at 80° C. The curve went to 4 000 hours and showed a few fluctuations (the measurements were not corrected to constant temperature), but there was no visible drift over this period.

The figures refer to transistors made by rather special techniques for communication capital equipment. They involve the complete absence of any 'goo' inside the case, most scrupulous cleanliness, and other features borrowed from vacuum-tube technology. This raises the question of cost, for the processes are more elaborate than those usually employed, and the transistor structure itself must be designed to fit these processes. Dr. Ebers mentioned that the cost of the finished wafer could be expected to fall to, say, 15 cents in four or five years' time, after mechanization of the production. Thus one might expect the cost of the complete transistor to fall to 4s.–6s. in due course.

The assumption that mechanization will take place depends on various factors, including the fact that the processes, products, and requirements of the circuit engineers will become stabilized, and also that the demand will be adequate.

Mr. J. N. Barry: In connection with Paper No. 2722, I certainly agree with Mr. Franklin that the question of capital cost could have been added to the list of advantages and disadvantages in Section 1. At present, and probably for some time in the future, it is likely to appear in the disadvantages column.

In Section 1.2, the authors comment that the ratio of transistors to valves for the same duty is about 2 : 1, while Mr. Phillips suggests that it is of the order of 2½ : 1. The authors go on to say that, with its triode construction, the transistor tends to require fewer components per stage than the valve. But a little earlier in the paper they say that, even at low frequencies, transistor circuit design is complicated by the greater number of significant parameters, some being dependent on operating conditions. The difficulty can be surmounted by suitable design techniques and the application of feedback. Surely both of these would require additional components. There do not, in fact, appear to be many fewer components per stage in the transistor circuits than in their valve counterparts.

Surely some of the additional components may be quite expensive, e.g. additional coupling transformers. It would be possible to employ quite a number of resistors and small capacitors for the cost of one transformer, particularly if the transformer had to comply with a tight transmission specification.

The second point is that a fairly rosy future is predicted for newer types of transistor, such as the diffused-base and drift

types. Although these devices certainly offer advantages in operating frequency, most types at present are limited both in respect of current and voltage operating range. These matters should receive some comment when predicting the use of such types.

Mr. King mentions better reliability when circuits are specially designed (presumably again using additional components). This raises important questions: What is to be considered as the end of useful life in transistors? To what extent are the parameters going to be allowed to vary with life? I should be interested if the authors of Paper No. 2764 could comment on the life variations they can tolerate in the transistor parameters used in their circuits, and particularly the current gain factor.

Mr. D. J. Marks: In Paper No. 2764 the authors refer to a number of types of a.f. amplifier, but not to one for the transmission of music. I have in mind an amplifier of, say, 50 c/s–10 kc/s bandwidth, with low distortion, such as the Post Office might use in providing lines for the broadcasting authorities.

In planning the advance provision of programme-type line plant, if it is to be economic, various difficulties arise because the provision has to be part of a network primarily intended for telephony and not for music transmission. A typical difficulty is that it may, in the event, be desirable to insert music amplifiers in places where there is no repeater station. This is where a transistorized music amplifier will have distinct advantages. With its extremely simple power-supply requirements it can be installed, at short notice, in almost any convenient building, e.g. an unattended automatic exchange. Taking the idea further, a 'buried' design will perhaps be practicable.

There is a considerable field of usefulness for a transistorized music line amplifier. Are there any special difficulties in its design?

Mr. M. B. Williams: In Section 2.1.7 of Paper No. 2722 a transistor version of a carrier repeater for a 12-channel single-cable carrier system is described. It would, perhaps, have been more generally useful to have completely revised the system to take full advantage of the properties of transistor amplifiers. For general application to deloaded a.f. cables, low-gain low-power amplifiers suitable for installation in underground chambers and for power feeding at low voltages are required. It would probably be more economical and certainly simpler to use separate pairs for the two directions of transmission than to use the split-band system with directional filters in each repeater.

Mr. J. Reynolds: References have been made to the desirability of reducing the standard output of a.f. line amplifiers when transistors are introduced. This question arose at the Post Office some years ago, because, with the transistors then available, the required output could not be obtained. The Post Office has large quantities of valve equipment working at certain standard levels, and it was agreed that no change in standards should be made unless it became certain that transistor amplifiers could not meet them. Development was continued as more suitable transistors became available, and this has led to a simple design, meeting all requirements, which can therefore be introduced without change in practice. This facility is well worth the extra development.

The amplifier TB4 met all the specified requirements. To improve reliability, development was continued to see whether the use of electrolytic capacitors could be avoided, and the TC1 amplifier was developed using only one small paper capacitor.

The TC1 amplifier needs no decoupling provided that the power-supply impedance is less than about 2 ohms, but this could not always be guaranteed in practice. Hence some decoupling has been provided in the final design. At the expense of some degradation in frequency response, a 4 μF paper capacitor has been fitted, thus retaining the feature of no electrolytic capacitors.

In Section 7 of Paper No. 2764, where a figure of $1{\cdot}9 \times 10^6$ transistor hours is given, the number of transistors was about 200.

Mr. M. P. Johnson: In the immediate post-war years, British 12-circuit carrier telephone systems were elaborately protected against component and valve failure. Owing to foreign competition, present-day systems appear on one bay without valve duplication or standby facilities. All the equipment is in plug form, and spares are stored in a cupboard. Are we prepared to go one step further with transistor equipment, and perhaps visualize 60-circuit terminals with no built-in standby equipment?

Mr. C. E. E. Clinch: There are a few references in the papers to the maintenance advantages of transistors. It is unusual in a design engineer's paper to mention maintenance advantages, but in this case it is the main selling point for transistors.

I am a little surprised at some of the figures on life, and particularly Mr. Cannon's suggestion of a hundred years. I hope that the authors can give more definite figures.

With reference to the fault rate, we have dispensed with the valve, and now, according to Mr. Eckersley, we are left with the capacitor as the main cause of failure. To this I would add the mechanical devices such as plugs and sockets and potentiometers. Designers of transistor equipment seem to put the plugs and sockets, which tend to have a high fault rate, at the back of equipment.

We might agree that we get some maintenance advantage from transistors. The lower fault rate will be achieved largely by the use of lower-voltage working. Some time ago I made a study of fault rates per valve stage in an attempt to find a common denominator between various types of equipment. This showed that the stages using higher voltages were giving more faults than those using lower voltages, probably mostly owing to capacitor failures. The use of lower voltages in transistor equipment should therefore bring better fault rates. I am worried when the authors suggest that they can reduce the costs by making cheaper components. This often tends to increase the failure rate, and it may offset the advantages of lower-voltage operation.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Messrs. H. T. Prior, D. J. R. Chapman, and A. A. M. Whitehead (*in reply*): Several speakers have discussed the relative advantages of valve and transistor equipment. The important matter is the overall economic advantage of transistor equipment, which, as Mr. Franklin points out, arises largely from simplification of the power supply and reduced housing requirements. Another important factor is the great opportunities for simplification of construction, mentioned by Mr. Cannon.

Reduced maintenance costs must also have a strong influence on the introduction of transistor equipment, and we support Mr. Johnson's idea that, for 60 circuits, spares could be provided by plug-in units stored in a cupboard. Mr. Clinch states that it would be undesirable to take up all the reduction of component requirements, brought about by the low voltages of transistor circuits, in the form of cheapening, particularly with capacitors. We agree with this, but on the other hand, the component fault rate will be further reduced by the reduction in general heating and hot spots. The former is expected to be about 5° C in transistor equipment compared with 15° C in valve designs. Thus there should be room both for cheapening and improved reliability.

Waltz's data, quoted by Dr. Tillman, on the failure rates of

valves and transistors relate mainly to valves of very low power efficiency. Brewer,* giving life-test figures on 15 000 commercial valves, indicates a failure rate after the first 100 hours' running-in period of 12% per 1 000 hours, and even the 'special quality' valves which are generally too expensive for use in line-transmission equipment have failure rates greater than 1% per 1 000 hours.

Long life with consequent reduced maintenance will only be attained by the very closely controlled highly mechanized methods of manufacture mentioned by Mr. King, and the cost will be greatly reduced if the number of types in use is kept small. In reply to Mr. Thomas, we consider that, for frequencies up to about 100 kc/s, one basic type of transistor element giving 50–100 mW dissipation in a small case and 0·5–1·0 watt in a large case with very low thermal resistance gives the designer sufficient scope. In reply to Mr. Barry, transistors giving a dissipation of about 50 mW with cut-off frequencies of about 10 Mc/s, and which are therefore useful up to about 2 Mc/s, are already available. Advances beyond this can be expected in the next few years.

Dr. Pim and Mr. Williams both mention the desirability of exploiting the properties of transistors in new types of systems not economically attractive with valves. Sections 2.2 and 2.3 of the paper describe some systems of this type, and there is no doubt that many other new systems will be developed along these lines. On the other hand, the very large investment in conventional systems will ensure a demand for designs compatible with existing valve systems (see Section 2.1) for many years to come.

It is suggested in the paper that, with transistor equipment, the a.f. output from carrier terminals should be reduced from the present relative level of + 10 dB (sometimes +13 dB) to perhaps +7 dB. Mr. Cannon welcomes this outlook, but Mr. Reynolds points out that the level of +13 dB could be obtained and considers the facility worth while. While there is no technical problem in obtaining a level of +13 dB we feel that the savings which result from limiting the output level would be greater than the expenditure on amplifiers at the receiving end of occasional long a.f. extensions.

Mr. Franklin doubts whether it is necessary to use p.c.m. systems to increase the number of circuits and improve the performance on junction circuits, and suggests that f.d.m. systems could be used. We agree that there is a field for f.d.m. systems in this application, but consider that, on the cables used in many foreign countries, problems of crosstalk and noise can only be dealt with by complicating the f.d.m. systems, e.g. by adding compandors. Signalling conditions are also frequently complex and differ greatly from territory to territory. Both considerations indicate a possible field for pulse code modulation.

Mr. Phillips points out that the 300-circuit system does not need a regulator every 4 000 yd with buried cable. This is true, and the system described in the paper will cope with the temperature variations of overhead cable. With buried cable there is no need to provide regulators at every repeater.

Both Messrs. Thomas and Barry dislike the transformers which are necessary in the interstages of common-base amplifiers. However, we must not overlook the fact that the common-base circuit has very definite advantages over the common-emitter when a very stable gain is required. It is possible to obtain from cascaded common-base stages, without overall feedback, a performance comparable to that of the common-emitter circuit with overall feedback. Thus it is unnecessary to include the transformers in the feedback path. Moreover, since no overall feedback is necessary when using common-base stages, the cut-off frequency of the transistors can be lower since there are no problems of controlling the $\mu\beta$ characteristic at frequencies well outside the pass band. The design techniques which we have in mind to overcome the difficulties introduced by extra significant parameters do not involve the extra complications and components suggested by Mr. Barry. They represent a suitable conservatism of design, so that the effects of variations in those parameters are not apparent.

Both Drs. Pim and Tillman refer to the use of transistors in modulators and filters. The application to modulators is very attractive for the reasons given by Dr. Pim. However, the need for great economy in carrier power is not yet urgent since the transistor equipment at present in use is designed to utilize the existing carrier supply equipment previously installed for valve equipment. The next phase of development may well be along the lines suggested by Dr. Pim. The use of transistors in filters, however, is not so likely. Transistors make the use of conventional active devices so much more flexible that considerable economies can be effected in the filters, which makes it unnecessary to incorporate transistors in the filters themselves.

Messrs. H. G. Bassett, D. Thomson, and P. E. Greenaway (*in reply*): We would like to support Mr. Franklin's remarks on the importance of the negative-impedance repeater for short a.f. circuits. A negative-impedance repeater using transistors has been designed, and a field trial of about 100 amplifiers is being commenced. In reply to Mr. Cannon, we agree that a moderate amount of feedback can be applied over two common-circuit stages and a transformer, provided that sufficient care is taken with the transistor and transformer specifications to ensure an adequate stability margin under all conditions. However, it is difficult to see how a reduction in the size of the transformer can be made unless loop gain is sacrificed at low frequencies. Mr. Eckersley and Dr. Tillman mention electrolytic capacitors. Investigations at the Post Office Research Station suggest that specially constructed capacitors using high-purity materials should, if used in paralleled pairs, perform satisfactorily for perhaps 20 years. If they become available, it will no longer be necessary to endeavour to avoid electrolytic capacitors in transmission equipment. The field trial referred to in the paper was carried out on equipment containing about 200 transistors.

In reply to Mr. Thomas, the most economical supply voltage usually depends on the peak collector voltage permitted. In amplifiers having output transformers the best supply voltage is often a little greater than half the maximum collector voltage, so that present-day alloy transistors with maximum voltages of about 30 volts are well suited for operation from the standard Post Office 21-volt supply.

In reply to Mr. Barry, the TB4 and TC1 amplifiers are designed around transistor specifications which allow the current gain to vary between 40 and 100. In reply to Mr. Marks, a music amplifier using transistors has been developed and is on field trial, although it is not intended for 'buried' service. No special difficulties attributable to the use of transistors have been encountered in the design.

In reply to Dr. Tillman, the non-uniformity of the presentation of information arises because of the widely separated times at which the amplifiers were developed. Change of I_{cb0} has negligible effect on the gain of the TB4 and TC1 amplifiers; change of I_{cb0} causes a slight shift of the operating point, which is, in no case, sufficient to bring the overload point of the amplifier below +17 dBm.

* BREWER, R.: *British Communications and Electronics*, 1958, 5, p. 258,

621.374.32:681.142

The Institution of Electrical Engineers
Paper No. 2859 M
May 1959

A DESIGN FOR AN AUTOMATIC GRAPH PLOTTER

By M. P. ATKINSON, B.Eng., W. T. BANE, B.Sc., Associate Member, and D. L. A. BARBER, B.Sc.

(*The paper was first received 25th April, and in revised form 30th December*, 1958.)

SUMMARY

A design for an automatic graph plotter using digital techniques throughout and incorporating full scaling and zero-shift facilities is described. Alternative approaches to some sections of the design are discussed.

The equipment makes use of transistors and printed circuits for both arithmetic and logical operations and incorporates a mechanical design for the table movements which makes possible the use of small low-powered driving motors with a high rate of acceleration.

With a suitable input device, points close together can be plotted at the rate of three per second with a positional accuracy better than 0·01 in.

(1) INTRODUCTION

Several machines have been designed to plot graphs automatically from digital information. Those incorporating scaling and zero-shift facilities can be divided into two classes.

In the first, the digital input information is converted into an analogue voltage on which operations are carried out to produce scaling and zero shift. The resulting signals are fed to analogue servo mechanisms which position a plotting head. In the second class, the positioning servo mechanisms operate directly from digital information, but scaling and zero shifts are performed using analogue methods (mechanical linkages).

The design to be described uses a digital system throughout. The use of digital techniques for scaling and zero shift makes the system less liable to accuracy limitation than its analogue counterparts. Because of the increase in complexity of the circuits, however, the system is an economic proposition only when use is made of transistors and printed-circuit techniques. The development was used to test the use of transistors both for arithmetic and logical operations in data-handling equipment.

(2) SPECIFICATION

The graph plotter was designed to the following specification:

(*a*) It should accept co-ordinates in the range 0–9 999 in x and y.
(*b*) The plotting area is to be 10 in square.
(*c*) The overall accuracy of plotting is to be 0·01 in in each co-ordinate, including scaling and zero-setting errors.
(*d*) The system should be adaptable to work with co-ordinates stored in either decimal or binary notation.

The system which was finally designed to meet the specification makes use of two digital accumulators (one for each axis), which are arranged to store the difference between the co-ordinates of the point to be plotted and those representing the present position of the plotting head.

Movement of the plotting head along either axis causes a series of pulses to be generated at equal intervals of movement along the axis. Each pulse causes a number (the scaling factor) to be subtracted from the appropriate accumulator, movement along an axis being stopped when the accumulator registers zero. When both accumulators register zero, the point is plotted. The co-ordinates of this point are then subtracted from the accumulators, and those of the new point are added in. The accumulators then contain numbers representing the distance to be moved to the next point, and the sequence is repeated.

(3) DESIGN CONSIDERATIONS AND DECISIONS

In order to plot a point automatically it is necessary to position a plotting head so that its distance from a preset origin, measured along the co-ordinate axes of the table, is proportional to the co-ordinates of the point. It is necessary either to convert the co-ordinates directly into linear movement on the table or to use a closed-loop system in which the linear movement along the axes is converted into a form which can be compared with the input, and the plotting head caused to move in order to minimize the difference between them. The digital plotting table to be described employs the latter system, and requires that the linear movement of the table be converted into digital form.

In order to carry out this conversion, the length to be digitalized is divided into equal small intervals. Movement within one of these intervals does not affect the digital reading of the position, which is changed by 1 each time movement occurs from one interval to the next. One system which has been employed in plotting equipments involves the use of a coded plate and appropriate reading device. For each interval along the plate a unique combination of code elements is set up by the reading device, and this combination represents the distance of the reading head from a fixed origin ('0' on the coded plate). The requirement of zero shift and scaling is not easy to fulfil with this arrangement, although complicated mechanical linkages, i.e. analogue devices, between the two reading heads and plotting head have been devised, which give limited scaling and shift facilities.

In view of these difficulties it was decided to use a counting system for digitalizing the movement of the plotting head. In such a system, movement of the plotting head along either axis causes a series of electrical pulses to be generated at fixed intervals of distance along the axis. These pulses are counted in a digital accumulator, which therefore contains at any instant a digital measurement of the distance of the plotting head from a preset origin. This origin can be set initially by moving the plotting head to the desired position on the table and setting the accumulators to zero, and the requirement of zero shift can easily be met.

At the outset it was decided that the requirement for scaling (i.e. that any range of numbers can be represented by any desired movement along the axes) should, if possible, be accomplished digitally. This can be done either by multiplying the input number by an appropriate scaling factor or by multiplying the fedback number by the reciprocal of the scaling factor. The first method, when examined in detail, involves fairly complex electronic equipment if it is to be accomplished in reasonable time, since it may be necessary to multiply a four-decimal digit number by another number of three decimal digits. The use of the counting system for digitalizing the table movement allows the second method to be used simply by adding into the accumulator the reciprocal of the scale factor each time a counting pulse

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is an official communication from the National Physical Laboratory.

is received. This method of scaling does, however, introduce complications in the electronic equipment used for detecting the 'balance point' unless some degeneration in accuracy is accepted. In order to appreciate this fact, it is necessary to consider the general principles governing servo-mechanism designs.

In order to control the drive along an axis of the table, it must be possible to decide whether the number held in the accumulator is greater than, less than, or equal to the desired number. This involves comparing and balancing each digit stage in turn, starting with the highest significant digit. An alternative system, which was used in this equipment, involves first adding into the accumulator the number representing the distance which the plotting head is required to move. The plotting head is now caused to move in a direction such that the counting pulses produced by the movement make the number in the accumulator tend towards zero, which now represents the required position of the plotting head. Thus the direction of movement is controlled only by the sign of the number held in the accumulator, and 'no error' between the actual and desired positions of the plotting head is always indicated by '0' in the accumulator.

The use of the system simplifies the problem of control of the plotting head, but it does introduce one extra difficulty. The number in the accumulator must always represent the distance to be moved. It must therefore be the difference between the new number to be plotted and that representing the present position of the table. After plotting any point, the number representing the present position of the table (i.e. the number which has just been plotted) must be subtracted from the next number to be plotted.

One advantage of a digital system is that its positional accuracy can be made equal to one count interval. When scaling is carried out by counting into the accumulator some fraction of '1' for each counting pulse produced, however, this accuracy becomes difficult to maintain. With scaling factor x, it is necessary to determine when the accumulator holds a number between 0 and x in order to ensure that the position of the plotting head is within one count of the desired position. Thus to detect the 'zero balance' position of the system, each stage of the accumulator below the binary or decimal point must be examined, and the 'zero range' (0 to x) will vary with the scaling factor. Clearly, if scaling is introduced by multiplying the input number by the appropriate factor and then counting out '1' for each pulse produced, this difficulty is not present.

With the feedback scaling system, however, if we count down a number in the range $\frac{1}{2}$ to 1 from the accumulator for each feedback pulse, then, by defining '0' as the position where all integer stages of the accumulator register zero, we cannot be more than two count intervals in error, and by setting the 'zero' of the counter to be 0·5, this two-count error can be made into ± 1 count.

In the design of the plotting table, it was decided to impose these limits on the feedback number. In order to obtain scaling factors between 0·1 and 0·5 it was decided to multiply the input number by 2, 4 or 8 during the process of adding the number into the accumulator. Since the accumulator works on the binary system, this multiplication merely involved the provision of a limited shift facility. The simplicity of application of this system was one reason why a binary accumulator was used. With a decimal system, where multiplication by 2, 4 and 8 would have presented a much greater problem, it is probable that some effort to detect the state of the first stage below the decimal point would have been made. 'Zero balance' would then have been defined as the position where all integer stages of the accumulator were zero and the stage below the decimal point held a number less than or equal to the highest decimal digit of the scaling factor.

The error due to the system as used in the design (excluding mechanical errors), will be less than ± 2 count intervals. This figure is deduced as follows. There is an inherent error of up to ± 1 count in the scaling system. There is an indeterminacy in the position of the table at any time of $\pm\frac{1}{2}$ count (the table position is only indicated by the counter, and movement of less than one count interval is not registered). Thus we can have a zero error of $\pm\frac{1}{2}$ count and a positional error of $\pm\frac{1}{2}$ count. In the worst case, with all these errors a maximum and of the same sign, we can have an error of $\frac{1}{2} + \frac{1}{2} + 1 = 2$ counts.

The desired accuracy of plotting was laid down as $\frac{1}{100}$ in. Thus it is necessary to produce counting pulses at intervals of $\frac{1}{200}$ in along the table axes. For the 10 in square table designed this means that full-scale movement on either axis must produce 2 000 counts.

These considerations led to the following requirements for the binary accumulators:

(*a*) The accumulator should be capable of recording numbers in the range −2 000 to +2 000. Thus there must be 12 binary stages above the binary point.

(*b*) To allow scaling factors of up to three decimal digits, it should be capable of adding or subtracting in parallel up to ten binary digits below the binary point, making a total of 22 binary stages in all.

(*c*) It must be possible to decide the sign of the number held in the accumulator at any time.

(*d*) It must be possible to detect the condition that all stages above the binary point hold zero.

(4) THE SERVO MECHANISM AND MECHANICAL DESIGN

The complete design is illustrated in Fig. 1. This Section gives details of the design and the reasons for its adoption.

With a normal analogue d.c. servo mechanism, use is made of the magnitude of the error between the actual and desired positions of the controlled moving member to control the speed of movement. If this principle had been used in the present system, it would have been necessary to convert the digital measure of error contained in the accumulators into a form suitable for controlling the motor drives to the two axes. This would have been further complicated by the use of the feedback system of scaling, which means that the number contained in the digital accumulator does not bear a constant relationship to the positional error.

In view of this difficulty, coupled with the fact that small overshoots before reaching a settled position could be tolerated, it was decided to design the servo mechanism with driving motors having only three possible states: full forward or reverse power controlled by the sign of the number in the accumulator, and no power if this number is zero. In the design chosen the motor drive is arrested by an electromagnetic brake, which operates for a limited time on receipt of a zero signal. If the plotting head overshoots and comes to rest outside the zero range, the accumulator will indicate this fact, the motor power is reversed and the braking operation is repeated. Only when the accumulator registers a steady zero position is the motor power switched off.

It is now necessary to consider the method of producing counting pulses. This was done by a device already described.* The member whose movement is to be measured is provided with a scale graduated with alternate opaque and transparent segments at equal intervals along the direction of movement. A beam of light is directed through the scale on to two photo-transistors positioned up to half a pitch interval apart along the axis of movement. When movement occurs, each photo-transistor in turn is illuminated and then not illuminated. These photo-transistors are connected into a circuit which produces a pulse

* BANE, W. T., and BARBER, D. L. A.: 'A Simple Direction-Sensitive Photo-transistor Circuit for use in Optical Pulse-Counting Systems', *Journal of Scientific Instruments*, 1956, 33, p. 483.

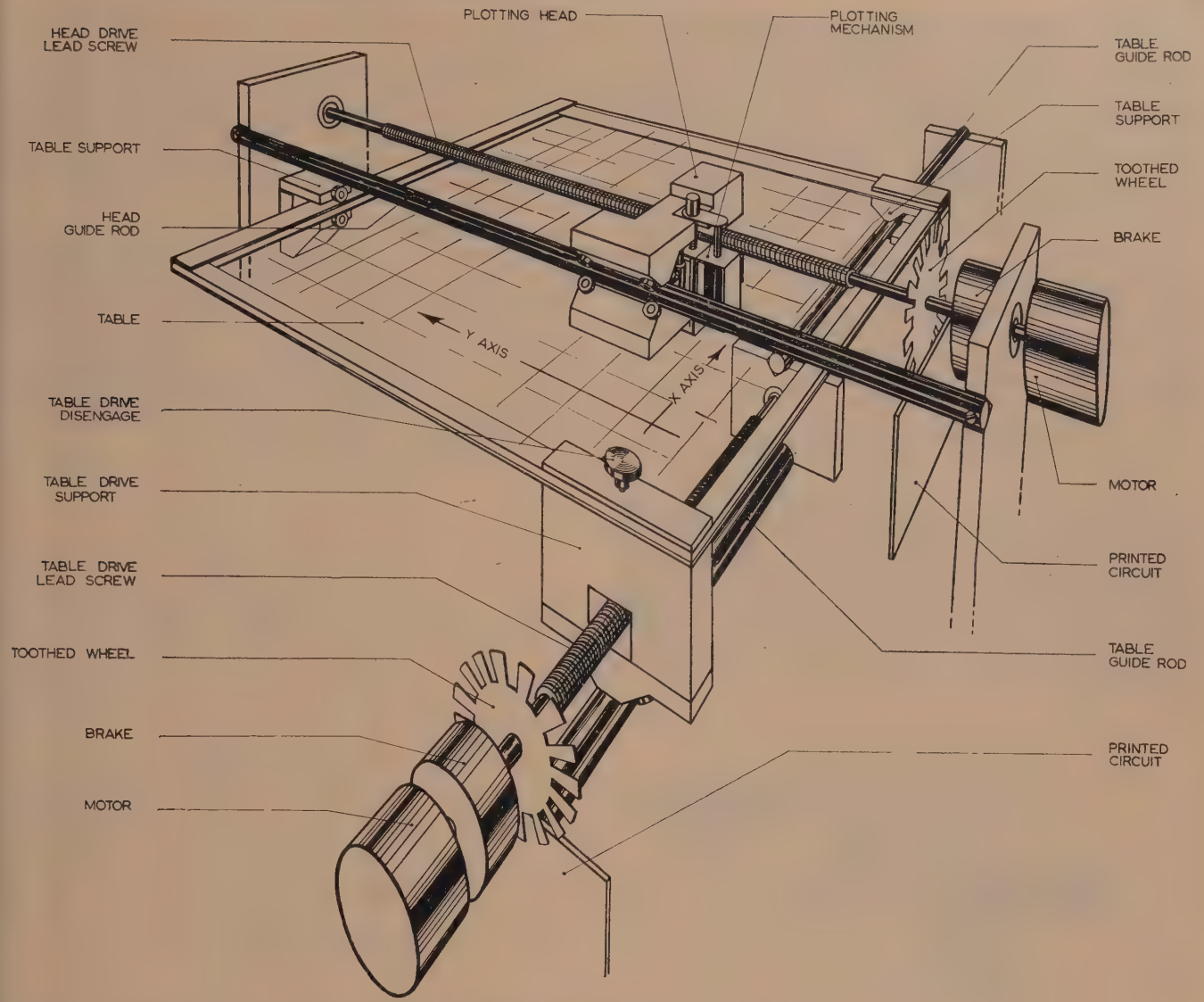


Fig. 1.—Mechanical arrangement of the plotting table.

on one of two lines, depending on the direction of movement, as each complete segment passes the light beam.

For the plotting table, these pulses are required for each 0·005 in of movement along an axis. The use of a linear scale on each axis would involve positioning the photo-transistors to an accuracy of 0·001 in, and owing to the variation in position of their sensitive elements, would involve some trial and error process. This fact, together with considerations of the amount of light which the scale would transmit through an 0·002 in transparent segment, made it probable that some optical magnification would have to be used within the measuring system. Furthermore the measuring scale must be 10 in long. If the pulses are to be produced by moving the scale past a fixed measuring system, each axis must be 20 in long. Alternatively, the measuring system itself must move past a fixed scale, involving movement of the optical system and electrical wiring.

It was therefore decided to mount the scale on the motor shaft. By this means, the scale becomes circular and can be magnified simply by increasing the radius at which the scale is marked. This, of course, means that the rotational movement of the shaft must be converted into linear movement to a much higher accuracy than would be the case if direct linear measurement were made along the axis of movement. The desired accuracy is easily obtainable in a lead-screw, and it was decided to use this type of drive.

The lead-screws used in the plotting table were cut with ten threads per inch. To obtain the desired 'count interval' of 0·005 in meant that each revolution of the lead-screw had to produce 20 counts. The lead-screw is therefore directly coupled to the motor shaft, on which is mounted a disc whose circumference is divided into 20 equally spaced slots. The disc is 3 in in diameter, which means that the pitch of the slots along the circumference is about $\frac{1}{2}$ in. The actual slots are 0·2 in wide. The tolerance on position of the photo-transistors is therefore relatively large.

Thus each axis must be equipped with a motor directly coupled to a lead-screw and an electromagnetic brake. The motor shaft carries a toothed disc, and the associated lamp and photo-transistors must be mounted in their appropriate positions to 'count' the passage of slots on the disc.

The design of the two axes can now be discussed. Along the y-axis is a gantry which carries all the driving equipment. The

plotting head is supported by this gantry and driven along it by the lead-screw. Movement along the x-axis is obtained by moving the whole table top underneath the gantry, which is fixed. This does, of course, mean that in the x-direction the equipment must be more than 20 in long (to allow a movement of 10 in of a table top which is 10 in long). However, in this design, the motors, brakes, etc., associated with each axis are fixed. If, instead, the gantry carrying the plotting head had been moved over a fixed table for the x-axis drive, all the y-axis driving and measuring equipment would have had to be movable, together with the associated wiring. Since the gantry would still require a supporting base (about 6 in long) in the x-direction, the saving in space would not have been very great.

The opportunity was taken to consider means of obtaining a high rate of plotting for points closely spaced. For this, a high acceleration is necessary. The acceleration of the motors is proportional to

$$\frac{\text{(Motor torque)} - \text{(Friction torque)}}{\text{Inertia}}$$

The inertia depends on the lead-screw diameter, the size of motor and brake, and the weight of the moving masses. If the motor size is to be kept down, it is necessary to keep friction torque down to a minimum. The great proportion of the friction torque is likely to occur in the means for converting the rotational movement of the lead-screw into linear movement along the axes, if the normal screw and threaded-nut arrangement is used. This problem was overcome by using a type of drive shown in Fig. 2.

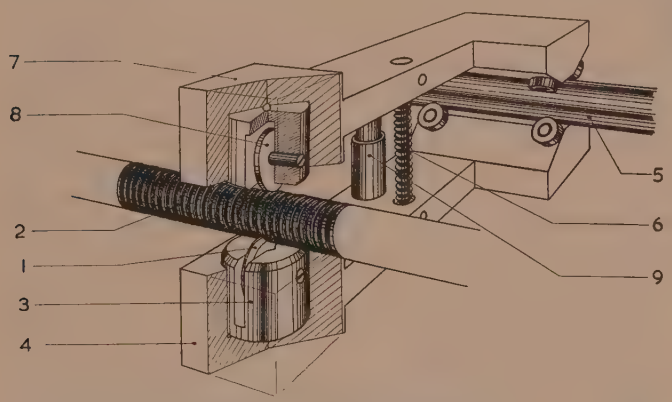


Fig. 2.—y-axis driving system.

A shaped wheel, 1, meshes with the lead-screw, 2, and acts as the 'nut'. This wheel is free to rotate on a spindle supported in a brass cylinder, 3, which is free to rotate about its own axis and allows the plane of the wheel to take up the angle of the lead-screw thread. This arrangement is mounted in one end of the block, 4, the other end of which is constrained to move along the guide rod, 5, by means of four rollers. The whole system is mounted below the centre-line of the lead-screw and guide rod, and is supported through the spring, 6, by a similar arrangement, 7, mounted above this centre-line. In the upper block, the roller, 8, replaces the shaped wheel of the lower block and rests on the flat tops of the lead-screw threads.

The whole of the driving force to this carriage is transmitted through the wheel, 1, which rotates with the lead-screw. Because of this rotation, sliding friction between the lead-screw and 'nut' is almost completely eliminated, and, in fact, is so small that it is possible to back-drive the lead-screw. The drive is transmitted from the lower to the upper block via the pillar, 9, which is attached to block 4 and fits loosely in a cylindrical hole in block 7.

With friction torque thus reduced to negligible proportions, the main limitation on acceleration and deceleration characteristics become the torque/inertia ratios of the motor and brake systems. It was found that, by using a lead-screw $\frac{1}{2}$ in in diameter and keeping the weight of the moving parts down to about 4 lb, a motor developing a torque of 5 oz-in could accelerate to full speed in $\frac{1}{4}$ sec.

An electromagnetic brake which develops a full torque of 16 lb-in is used to stop the motion. When the manufacturers' recommended circuit was used to operate the brake, however, it was found that full torque was not developed until 30 millisec after applying the operating voltage, and the rate of increase of the torque was linear over this period. Because of the high acceleration characteristics of the motor drive, this meant that several overshoots would occur before the plotting head reached a settled position. The circuit shown in Fig. 3 was therefore used to operate the brake.

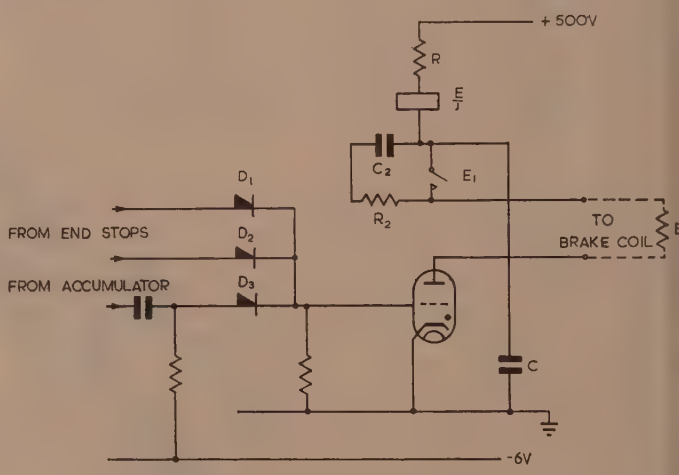


Fig. 3.—Brake operating circuit.

A pulse, from an accumulator reaching zero or an end stop operating, fires a thyratron, which discharges a condenser C through the electromagnetic brake B. The condenser-recharging resistor R is in series with a relay E which operates after a delay and resets the circuit. Diodes D_1, D_2 and D_3 combine the input signal and also protect the driving transistors from damage by the positive grid pulses produced when the thyratron strikes. Using this circuit, full braking torque is developed in 2 millisec.

The waveforms of Fig. 4 illustrate the operation of the circuit.

The circuit used to operate the driving motors is shown in Fig. 5. Transistor-operated relays A and B are used to supply power to a split-field induction motor. A or B is energized, depending upon the sign of the number in the accumulator, to give appropriate in-phase and quadrature signals to the motor windings W_1 and W_2.

The operation of an end stop releases the appropriate relay, and provides a positive pulse to fire the brake thyratron.

Two thyratrons connected in a monostable circuit (Fig. 6) are used to operate the plot stamping head. Thyratron T_2 will normally be conducting. When an input pulse causes T_1 to strike, T_2 is extinguished by the pulse through C_1, and the plot coil is supplied with a large negative pulse which operates the stamping head. The voltage across the coil decays towards zero until T_2 strikes and resets the circuit.

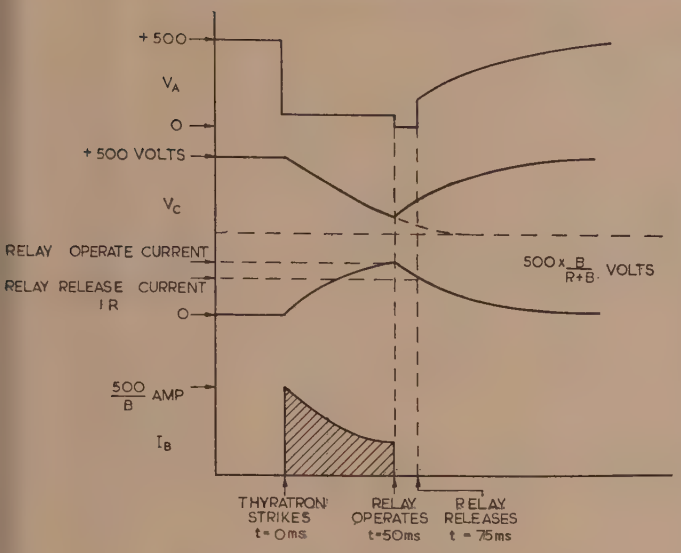


Fig. 4.—Brake-circuit waveforms.

(5) THE ACCUMULATORS

The co-ordinates of the point to be plotted must be added into the appropriate accumulators before plotting. After plotting, these co-ordinates must be subtracted before co-ordinates of the next point are set in. The system described was designed to accept information which is stored in decimal form. Thus it is necessary to convert the information into binary form before or during the process of addition into the accumulators.

A four-decimal digit number can be considered to be comprised of a '100's digit, a '10's digit, a units digit and a '0·1's digit, e.g. 765·4. The binary equivalents of 100, 10, 1 and 1/10 are as follows:

Decimal	Binary
100	1100100.
10	1010.
1	1.
0·1	0·0001100110 . . .

Thus to add the decimal number 111·1 into the binary accumulator we must add 1 to the ten binary stages listed above. Since no binary stage is included in more than one decimal equivalent, only one driving pulse will be required to perform this addition, the pulse being routed to all ten stages. To add 999·9 we add 111·1 nine times, which requires nine pulses in all.

In the system used, ten lines are each connected to all four stages of the decimal store and each stage has one output to the appropriate binary stages of the accumulator. For instance, the '100's stage has its output connected to the 2^6, 2^5 and 2^3 stages of the accumulator. If the ten input lines are numbered 0 to 9, the presence of a digit in any stage of the accumulator connects the output of the stage to the appropriate input line (i.e. 5 in the '100's stage connects line 5 to the output of the '100's stage and hence to the three equivalent binary stages of the accumulator).

When addition is required the ten lines 0 to 9 pass 0 to 9 pulses, respectively, to all four decimal stages. Line 5, for instance, passes only five pulses during one addition cycle. The storage elements then route the desired number of pulses to the appropriate binary stages of the accumulator.

There is a certain minimum time (in this case, 25 microsec) which must be allowed between pulses to any stage in order that it shall work effectively. Since several stages are being operated by the same input pulse, and any stage will produce a 'carry' pulse on changing state from '1' to '0', it is necessary to delay this pulse by at least 25 microsec before passing it to the next higher stage. The upper limit on this delay is governed by the condition that a 'carry' pulse generated by the lowest input stage which causes each higher stage to change in turn must reach the highest input stage before the next input pulse is generated. In our case, the limiting stages are $\frac{1}{5}\frac{1}{12}$ and 64, and there are 15 delays between these two stages. If counting pulses are generated at a frequency of 1 kc/s the delay time must be less than 65 microsec. In order to simplify the design all stages are identical and have two inputs, one of which contains a delay.

The pulses are produced by pulse generator (Fig. 7). This

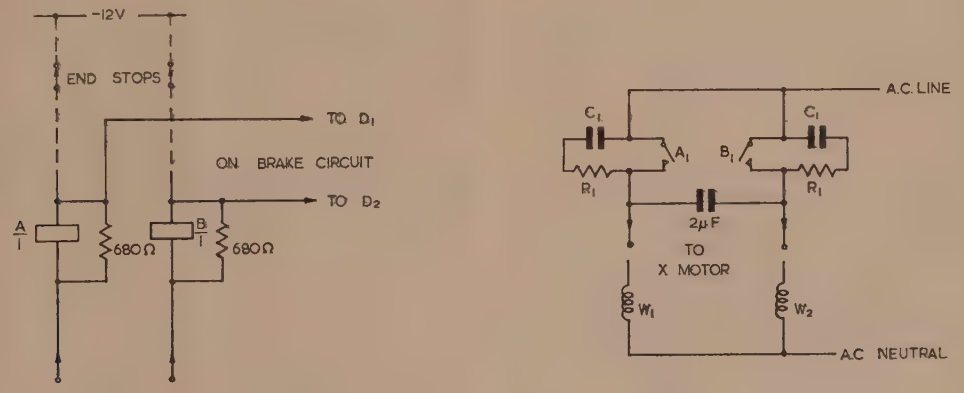


Fig. 5.—Motor operating circuit.

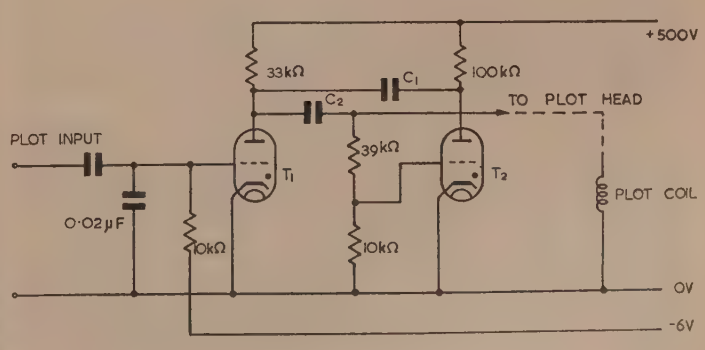


Fig. 6.—Stamping-head operating circuit.

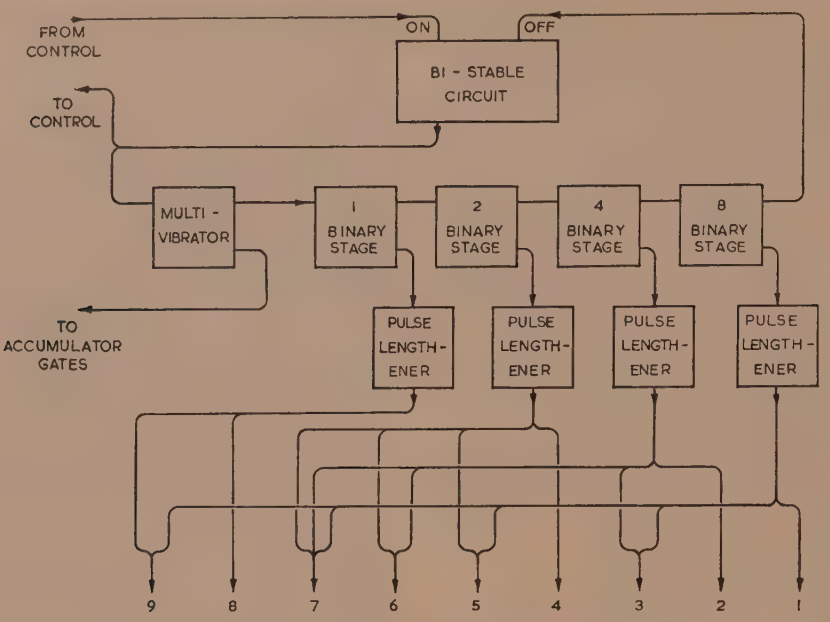


Fig. 7.—Schematic of pulse generator.

consists basically of a four-stage binary counter and a multivibrator. A signal from the control unit of the machine starts the multivibrator, successive pulses from which are 'counted' by the counter, which transmits a 'stop' signal to the multivibrator after 16 pulses have been received. During this cycle, stage 1 of the counter changes state from '0' to '1' and back a total of eight times. Similarly stage 2 changes four times, stage 4 changes twice and stage 8 once. It may be noted that no two stages can change from '0' to '1' simultaneously. Four lines, one for each stage, are connected to the stage in such a way that they transmit a pulse each time that stage changes from '0' to '1'. These pulses are then lengthened and added in the resistor-transistor network to produce the ten pulse lines required.

In practice, these long pulses are not used directly to add into the accumulators, but are used to 'gate' sharp pulses from the multivibrator, which are then fed to the accumulators. This means that long connecting leads between the machine and the digital information source will not seriously affect the form of pulse used by the accumulators. The circuit used for these gates is shown in Fig. 8.

The system requires that a number to be plotted shall be both added to and subtracted from the accumulators in the plotting sequence. The system of subtracting a number by adding its complement is not suitable for this application. For example, to subtract 765·4 would involve either:

(*a*) Adding the binary complement of 100 seven times, of 10 six times, etc. But, since several binary stages are involved in the complements for all four decimal stages, the addition would have to be carried out one decimal stage at a time.

(*b*) Adding the '9's complement of the decimal number (i.e. 234·5) and then adding a fixed correction number to convert this to the binary complement. The '9's complement is easily obtained by operating a relay whenever a negative sign is present. This would cause the connections from the pulse generator to be interchanged so that line 9 received no pulses, line 8 received one pulse and so on. However, if the x and y co-ordinates are of opposite sign, one must be added when the other is subtracted. Independent connections from the pulse generator to the x and y stores are therefore necessary, which places restrictions on some types of storage system from which the co-ordinates might be obtained.

It was thus decided to design the accumulators to subtract numbers when set up to do so. When adding a number, a 'carry' must be generated whenever a stage changes from '1' to '0'. In subtraction a 'borrow' is generated on changing from '0' to '1'. The system is so arranged that, in normal operation, each stage (when it changes from '1' to '0') passes a pulse to the next (higher) stage, but when set to subtract, this pulse is passed on changing from '0' to '1'.

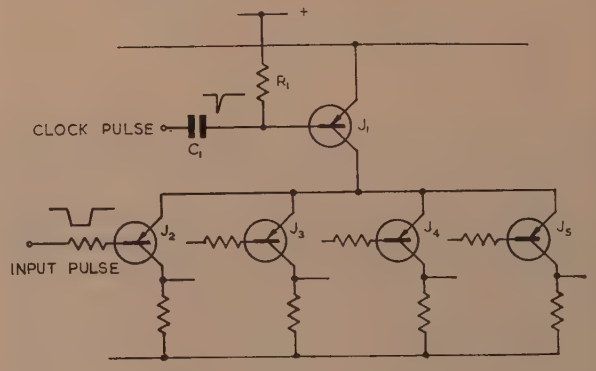


Fig. 8.—Input-pulse gate circuit.

The feedback system provides pulses on one of two lines, depending on the direction of motion of the driving system. Pulses on one line must cause the preset scaling number to be added into the accumulator, whilst those on the other line must subtract the scaling number. The two pulse lines are connected in such a way that a pulse on the appropriate line sets the accumulators to add or subtract. The pulse lines are then combined and fed via a delay to the scaling switches and hence to chosen stages of the accumulators. Because of the delays in the 'carry' lines, this scaling number can be added in parallel.

The circuit used for a binary counting stage is shown in Fig. 9.

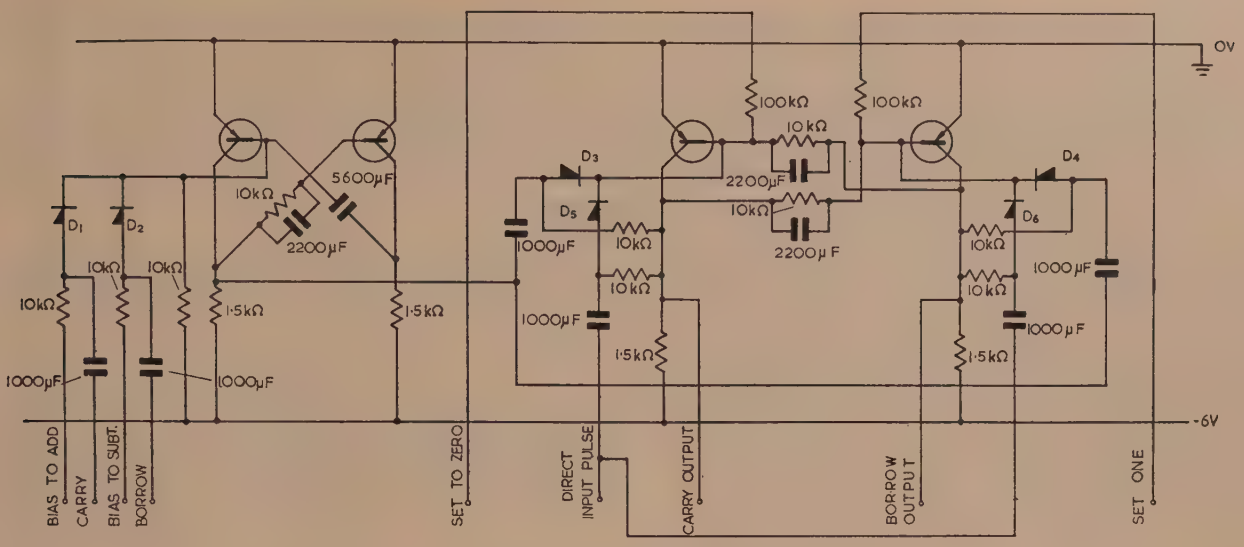


Fig. 9.—Binary counter circuit.

The 'carry' (add) or 'borrow' (subtract) pulse from the previous stage is selected by biasing the appropriate diode gate D_1 or D_2, and it triggers a monostable delay circuit. The output pulse from this is applied to an Eccles–Jordan bistable circuit through a pair of diodes D_3 and D_4 arranged to form an input pulse-steering circuit. Another similar pulse-steering circuit with diodes D_5 and D_6 enables direct input pulses to be fed to the binary stage.

A miniature tuning indicator provides a means for visual detection of the state of the counter stage.

The circuit for detecting the accumulator zero is shown in Fig. 10. This is a multiple input gate. The 12 inputs (one from each stage above the binary point) connect to the base of J_1 via 100-kilohm resistors, in such a way that the input from any stage comes from a bottomed collector when that stage registers zero. With any one stage in the 'non-zero' condition, the equivalent input line is connected to a negative potential, which is sufficient to cause J_1 to be bottomed. Thus J_2 is almost cut off and the output potential is negative. Only with all stages at zero does J_1 cut-off, and when this occurs J_2 bottoms and the output rises to near earth potential.

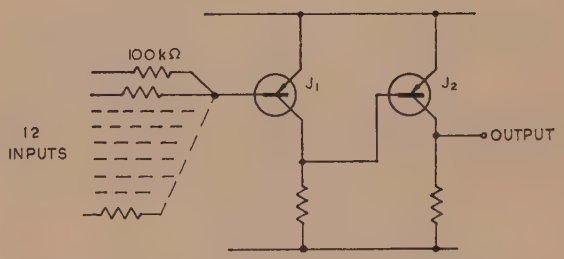


Fig. 10.—Accumulator zero detecting circuit.

(6) THE SEQUENCE CONTROL SYSTEM

In order to plot a point the following sequence of operations must be carried out:

(*a*) Add the co-ordinates of the point into the accumulators.
(*b*) Supply power to the motors, controlling the directions of motion according to the signs of the numbers held by the accumulators.
(*c*) As each accumulator reaches '0' apply the appropriate brake.

If there is an overshoot, reverse the motor, return to zero and apply the brake again. When settled at zero, remove the motor power.

(*d*) When both axes are settled at zero and the motor power has been removed, subtract the co-ordinates of the point just plotted from the accumulators.
(*e*) Signal the digital reading equipment to read the next point, and wait for a signal from this equipment before restarting the sequence.

The above operations must be carried out in the order stated, without overlapping. The system adopted in this machine is that each of the above operations is regarded as a sub-sequence, which is started by a signal from the central control, to which it transmits a signal on completion of the operation. The central control system is then simply a 'counter'.

Before plotting a graph, it is necessary to set the desired zero of the plotting head and the correct scale factor. The plotting head must first be moved to the desired zero position on the table. The scale factor must then be set so that a given number in each co-ordinate will correspond to a desired movement on the appropriate axis. When the plot is to be made on plain paper this setting-up is straightforward, since adjustment of the scaling switches will enable any range of numbers to be plotted as any desired movement to an accuracy of $\frac{1}{100}$ in. This scale factor can be calculated and set up on switches.

When plotting on pre-printed graph paper, however, some trial-and-error adjustment of the scale switches may be necessary. By measuring the required range, the scale factor can be calculated, but some slight modification to this number may be necessary, owing to errors in measurement. This requirement means that additional facilities are necessary in the control system.

The desired numerical range for each co-ordinate can be set up on switches and the plotting head set to the zero point on the graph paper. The system must then, on instruction, plot the point indicated by the range switches. If this is not the desired point on the graph paper, the scale factor must be adjusted and the operation repeated. In order to facilitate this, the sequence is arranged to return the plotting head to the zero position after plotting the full-scale point.

The range-setting operation is carried out by modifying

the main plotting sequence. A three-position switch, marked 'normal', 'set minimum' and 'set range', is incorporated in the machine. On setting this switch to 'set range' the accumulators are cleared. On operating a 'start' pushbutton, operations (*a*) to (*d*) of the main sequence are carried out. Operation (*e*) is replaced by a repeat of operations (*b*), (*c*), (*d*) and (*e*).

It may be necessary to make the origin of a given graph represent a number other than zero in either co-ordinate. Hence, at any position, the number representing the origin must be subtracted from the accumulators in the initial setting-up procedure, after the range has been set. The 'minimum' co-ordinates can be set up on switches, and with the three-position switch rotated to 'set minimum', depressing the 'start' pushbutton causes operation (*a*) of the main sequence to be carried out with the signs of the minimum co-ordinates inverted.

The whole of the control system has been transistorized, relays being used where necessary as 'slaves', to control the operations of motors and brakes, which require high voltages. The control circuit is shown schematically in Fig. 11.

(*a*), starts PG and resets BS_2. BS_2 now opens G_2, and the return signal from PG is directed to set BS_4, which operates the relay CS to clear the stores and resets BS_1. BS_4 is reset by a signal from the stores when they are clear.

The setting-up operations modify the sequence as described above. For setting the minimum number, BS_2 is not set up, so that gate G_2 remains open and the return from PG after operation (*a*) causes steps (*b*), (*c*) and (*d*) of the sequence to be omitted.

For 'set-range', the signal through G_2 is routed to BS_5, which restarts the motors. The return signal now resets BS_5, which sets up BS_4, and the sequence ends as before.

(7) PERFORMANCE

With the co-ordinates of each point to be plotted presented to the machine in parallel, a plotting rate of 3 points per second has been achieved for points close together (about $\frac{1}{4}$ in spacing). The average accuracy of plotting a point is about 0·007 in,

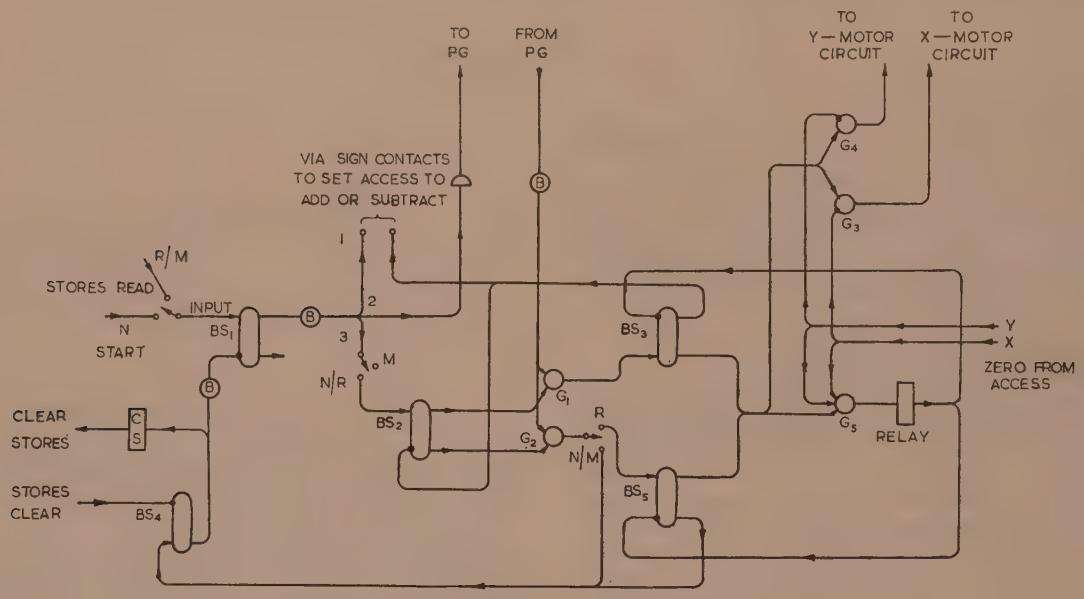


Fig. 11.—Schematic of control circuit.

N. Normal.
M. Minimum range.
R. Range.

Beginning element.

Input → (B) → Input differentiated.

Bistable circuit.

On → BS → Output when 'on'.
Off → BS → Output when 'off'.

The start signal sets up bistable circuit BS_1, the output pulse from which is connected to do the following:

(*a*) Set the accumulators to add or subtract according to the sign of the co-ordinates.
(*b*) Start the pulse generator PG after a delay.
(*c*) Set up BS_2.

The output from BS_2 opens the gate G_1, and the return signal from PG passes through G to set_1 up BS_3, which, in turn, sets up the X and Y motor circuits, via G_3 and G_4. These gates are closed when the appropriate accumulators are at zero. The '3' gate G_5 has a relay in circuit, which operates when both accumulators are at zero and BS_3 is in the 'start motors' condition. The time of operation of this relay ensures that the two axes are in a 'settled' zero condition. The relay then causes the plot solenoid to operate and resets BS_3, a pulse from which sets the accumulators to add or subtract again in the opposite sense to operation

although a point can be repeated with greater accuracy (about 0·003 in).

The use of printed circuits in the accumulators eased the problems of servicing and fault finding in the experimental stages of the design. The use of a pulse-counting system for digitizing the co-ordinate movement has proved particularly successful. Over a considerable period of time, the authors have not detected a pulse being missed.

(8) ACKNOWLEDGMENTS

The development was based on general principles proposed by Mr. R. H. Tizard, to whom the authors wish to acknowledge their indebtedness.

The work described was carried out as part of the research programme of the National Physical Laboratory, and is published by kind permission of the Director.

621.317.33: 621.319.42: 389.6

The Institution of Electrical Engineers
Paper No. 2887 M
May 1959
©

THE CYLINDRICAL CROSS-CAPACITOR AS A CALCULABLE STANDARD

By A. M. THOMPSON, B.Sc.

(*The paper was first received 27th November, 1958, and in revised form 2nd February, 1959.*)

SUMMARY

A type of cylindrical capacitor is described for which the capacitance per unit length may be computed with very great precision. It consists of a hollow conducting cylinder divided into four insulated sections by gaps parallel to its axis. The defined capacitance is the direct capacitance between pairs of internal opposing faces of the cylinder. If the two capacitances so obtained are equal, both are equal to $(\log_e 2)/4\pi^2$ e.s.u./cm whatever the shape of the cross-section of the cylinder. If the two cross-capacitances are not quite equal, the mean capacitance per unit length is very nearly equal to the above value. A measurement of the effective length of the capacitor is the only mechanical measurement required for the computation of the mean cross-capacitance.

The length may be defined by using guard electrodes, and techniques are described for eliminating uncertainties in the position of the guard gaps. An alternative is to measure the change of capacitance for a known movement of a terminating shield.

An experimental capacitor constructed from round bars is described. The computed capacitance (1 pF) agreed to 1 part in 10^5 with the value measured in terms of standards of resistance and frequency.

(1) INTRODUCTION

During recent years the uncertainty in the measured value of the velocity of light* has been reduced considerably. It is now considered to be known to about 1 part in 10^6, which is an order better than is claimed for the electrical units based on the calculation of inductance. This has led to an examination of the possibility of obtaining more precise values for the practical electrical units by the computation of a suitable capacitance. Since very precise mechanical measurements are necessary, a calculable capacitor must be constructed from those surfaces which can be generated with the required precision. Most consideration has been given to capacitors constructed from planes and cylinders of circular cross-section. Parallel-plate and concentric-cylinder capacitors are both possibilities. A variable-spacing parallel-plate capacitor was suggested by the author's colleague Mr. W. K. Clothier, and his preliminary work indicated that it should be possible to obtain an accuracy of 1 part in 10^6 but that it would be very difficult to improve on this.

From a consideration of cylindrical systems the author suggested a new type of cylindrical cross-capacitor.[1] The capacitance per unit length of any cylindrical capacitor is dependent only on the shape of the cross-section, i.e. on the ratios of dimensions, and since unity is the ratio most easily determined experimentally, the possibility of constructing a capacitor from a hollow conducting cylinder of square cross-section was examined. It was proposed that the four sides of the cylinder should be separated by insulating gaps and that the defined capacitance should be that between internal opposing faces of the cylinder. It was suggested that, if the asymmetry were not too great, the mean of the two cross-capacitances would give a very close approximation to the value obtained for a perfectly square cross-section.

Lampard investigated the cylindrical cross-capacitor of square cross-section theoretically[2] and showed that the capacitance per unit length in a vacuum would be $(\log_e 2)/4\pi^2$ e.s.u./cm and that the error in the mean value due to a small asymmetry would be of the second order. Other symmetrical cross-sections were suggested, and in each case the analysis gave the same capacitance per unit length. This led Lampard to propose and prove that any hollow conducting cylinder whose cross-section has one plane of symmetry could be divided into two pairs of similar electrodes such that the two cross-capacitances per unit length would both be equal to

$$(\log_e 2)/4\pi^2 = 0{\cdot}017\,557\,623 \text{ e.s.u./cm.}$$

If the velocity of light is taken as 299 792·50 km/s this is equivalent to 0·019 535 485 pF/cm.

With the symmetry condition established as the only requirement of the cross-section, the author suggested the form of capacitor shown in Fig. 1. The capacitor is constructed of four

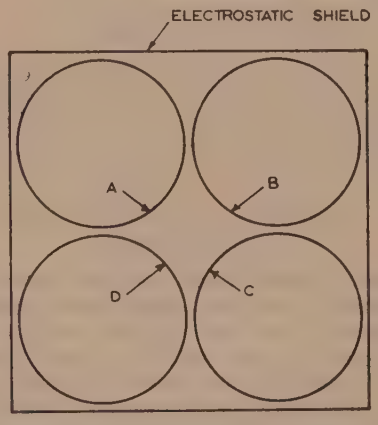


Fig. 1.—Cross-section of cylindrical cross-capacitor constructed from four bars.
$C_1 = C_{A-C}$.
$C_2 = C_{B-D}$.

similar cylindrical bars of circular cross-section. With this form the necessary gaps between the electrodes can be quite large without appreciable effect on the capacitance. An experimental model of a capacitor of this type was constructed and many measurements were made to determine its limitations as a calculable standard.

A recent re-examination of other possible cross-sections led to the conclusion that Lampard's condition of mechanical symmetry, although sufficient to give equal cross-capacitances, is not necessary. It may be shown that the only requirement for a calculable cross-capacitor is that its cross-capacitances be equal, and this condition may be obtained for any cross-section by moving one of the insulating gaps.

* A recent determination by K. D. Froome (*Nature*, 1958, **181**, p. 258) is 299 792·50 ± 0·10 km/sec.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Thompson is in the Division of Electrotechnology, National Standards Laboratory, Commonwealth Scientific and Industrial Research Organization, Australia.

(2) THEORY

A general relation connecting the two cross-capacitances may be obtained by conformal mapping of the cross-section on to a half-plane. This case was considered by Lampard and he showed [Reference 2, eqn. (75)] that the cross-capacitances per unit length in a vacuum are of the form

$$C_1 = \frac{1}{4\pi^2}\log_\varepsilon\left(1+\frac{a}{b}\right) \quad . \quad . \quad . \quad . \quad (1)$$

$$C_2 = \frac{1}{4\pi^2}\log_\varepsilon\left(1+\frac{b}{a}\right) \quad . \quad . \quad . \quad . \quad (2)$$

If $C_1 = C_2$

then $C_1 = C_2 = C_0$, where $C_0 = \frac{1}{4\pi^2}\log_\varepsilon 2$. . (3)

If the cross-capacitances are not quite equal and $C_1 - C_2 = \Delta C$, then, from eqns. (1) and (2),

$$\frac{C_1 + C_2}{2C_0} = 1 + \frac{\log_\varepsilon 2}{8}\left(\frac{\Delta C}{C_0}\right)^2 + \ldots \quad . \quad . \quad (4)$$

This indicates that, if the cross-capacitances differ by 1 part in 10^3, the error in the mean value due to ignoring the second-order term in eqn. (4) is less than 1 part in 10^7. The separate capacitances C_1 and C_2 may be computed from eqns. (1) and (2) for any value of C_1/C_2 if this ratio can be determined with sufficient accuracy. However, all practical considerations suggest that this ratio should be approximately unity, as in the case considered above.

Eqns. (1) and (2) may also be combined to give a general relation* valid for all cross-sections:

$$\exp - 4\pi^2 C_1 + \exp - 4\pi^2 C_2 = 1 \quad . \quad . \quad . \quad (5)$$

We have been concerned with the capacitances between internal surfaces of the cylinder. The same relation holds for the capacitances between external surfaces, but in practice these external capacitances are eliminated by an electrostatic shield.

(3) DESIGN CONSIDERATIONS

(3.1) Cross-Section

The most suitable cross-sections for calculable cross-capacitors are those in which the necessary gaps between the electrodes are re-entrant, since in this case the gaps may be quite large [Reference 2, eqn. (85)]. Approximate equality of the two cross-capacitances may be obtained either by the use of a symmetrical cross-section or by mechanical adjustment guided by capacitance measurements. Some of the many cross-sections which should be suitable for the construction of precise calculable cross-capacitors are shown in Fig. 2. The choice of cross-section may be influenced by the method used to define the length of the capacitor.

(3.2) Definition of Length

Two defined lengths are necessary, one for each of the cross-capacitances, and these should be approximately the same and occupy the same portion of the capacitor. The simplest method of defining the length of one of the cross-capacitances is to isolate a section of one of the two conductors concerned by means of insulating gaps and to measure the direct capacitance of the isolated section. This is the familiar guard-electrode technique. Its most serious limitation is that errors in sym-

* This relation was derived by Lampard for two special cases [Reference 2, eqns. (32) and (71)] and he also pointed out its generality in a C.S.I.R.O report dated March, 1957. A similar general relation has been established by L. J. van der Pauw for the corresponding 4-terminal resistances of a cylindrical resistor.[6]

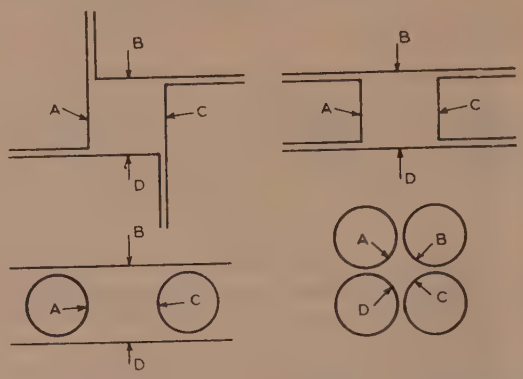


Fig. 2.—Possible cross-sections for a precise calculable cross-capacitor.
$C_1 = C_{A-C}$.
$C_2 = C_{B-D}$.

metry of the gap give rise to comparable errors in the electrical length. It is very difficult to obtain a precision of 10 microin with mechanically assembled gaps. Clothier has made gaps by evaporating a metal on to a polished insulating surface, the shadow of a stretched wire forming the gap. These were very satisfactory, and gaps of this type would almost certainly eliminate any difference between the mechanical and electrical lengths. The gaps can be made so narrow that the error due to their finite width is negligible.

An alternative to very precise gaps is to measure the effective electrical length of the assembled capacitor by determining the electrical centre of the gap with a suitable capacitance probe. The movement of the probe may be measured by coupling it to an interferometer. Even if the electrical centre of a guard gap is uncertain it is still possible to eliminate this uncertainty by using either the sum or difference of capacitances obtained by mechanical rearrangement of the capacitor. In the first case the length-determining conductor is made of two sections, each of which contains an insulating gap, as shown in Fig. 3. The

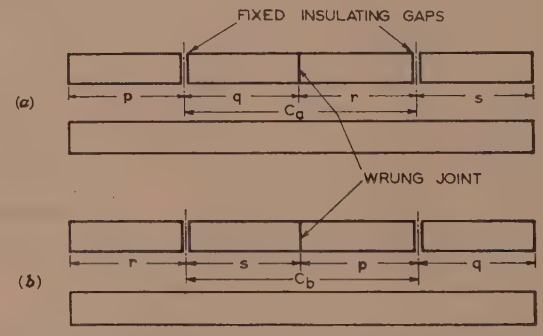


Fig. 3.—The use of two composite end-bars and two separate assemblies to eliminate uncertainties in the positions of the insulating gaps.
$C_a = C_0(q + r)$.
$C_b = C_0(s + p)$.

capacitor may be assembled, without in any way changing the geometry of the insulating gaps, by wringing the two sections together at either of their ends. The sum of the capacitances obtained for the two assemblies shown in Fig. 3 is that corresponding to the total length of the system. The gaps may be placed so that all capacitances are of the same nominal value.

The second technique is to have two separate pieces containing the insulating gaps as above, and to determine the change in capacitance as these are separated by an additional piece of known length. A capacitor of this type has been constructed

by the National Bureau of Standards,[3] and preliminary results indicate a reproducibility of 1 part in 10^6.

It is unnecessary to use guard gaps to define the length of a cylindrical capacitor. Conducting shields inserted into each end of the capacitor, as shown in Fig. 4, may be used to limit

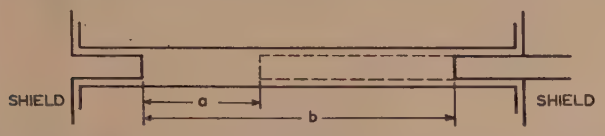


Fig. 4.—Defining the length of a cylindrical capacitor by conducting shields.

$$C_b - C_a = C_0 (b - a)$$

the effective length. The change of capacitance is measured for two different separations of the shields. Since the one shield defines the length of both cross-capacitances, only the one measurement of length is necessary. With this type of shield there is considerable distortion of the fields at the ends of the capacitor and there is an unknown difference between the length of an equivalent ideal capacitor and the mechanical separation of the shields. However, the change of capacitance for a change in the separation of the shields is exactly equivalent to that produced by the same change in length of an ideal capacitor. The field distortions due to the shields must be reproducible at the two separations, and the minimum separation of the shields must be such that there is no appreciable interaction between the two end-effects.

(4) EXPERIMENTAL CAPACITOR

A capacitor of the guarded-length type was assembled from a number of ground stainless-steel bars, two of which were constructed as precision end-bars with insulated end-pieces. The mode of assembly is illustrated in Fig. 5.

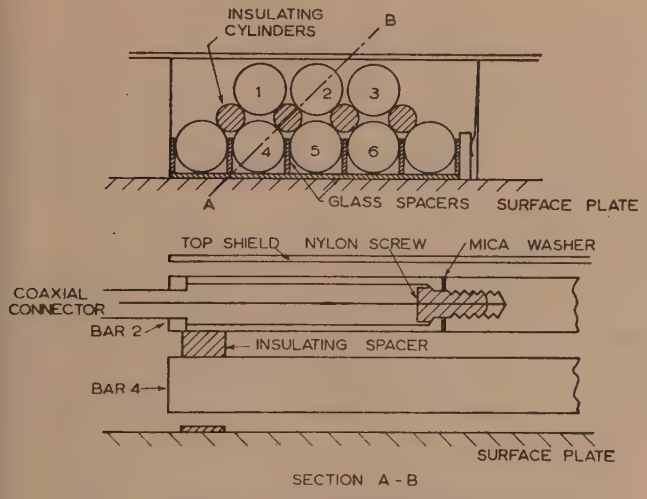


Fig. 5.—Mode of assembly of experimental cylindrical cross-capacitor.

Five 1 in-diameter bars about 18 in long were supported near their ends on glass strips on a surface plate. They were fixed parallel to each other by small glass spacers. Short insulating cylinders resting in the gaps between these bars supported the top three bars. This gave a symmetrical assembly, the capacitance asymmetry of which was always less than 1 part in 10^3. Bars 2 and 5 contained the guarded sections, which were just over 10 in long. The insulating spacers and supports were at the ends of the bars and well clear of the working section of the capacitor.

To determine the effect of the gaps between the bars, the direct capacitance between one of the 10 in guarded sections and a shield behind one of the gaps was measured. The results are shown in Fig. 6. These measurements show that gaps as

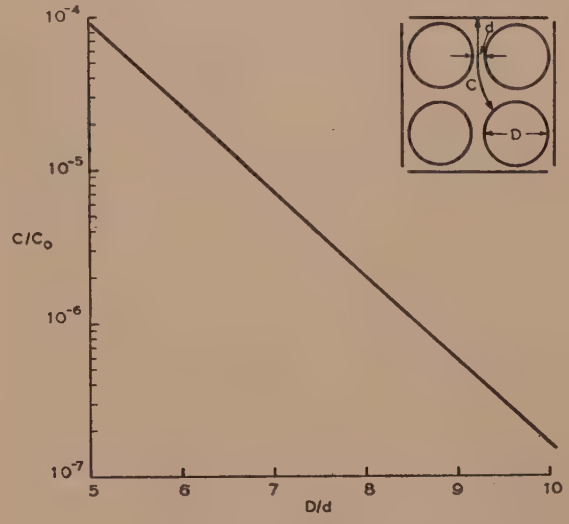


Fig. 6.—Capacitance through gap from one bar to shield.

large as one-tenth of the diameter of the bars cause an error of less than 2 parts in 10^7 in the calculated capacitance.

Tests show that the capacitor is insensitive to imperfections in the bars and the assembly, except where these alter the definition of the length of the capacitor. For example, a strip of tin-foil 0·0005 in thick and 2 in wide wrapped around one of the bars affected the mean capacitance by more than 1 part in 10^6 only when it approached the ends of the capacitor. Similar results were obtained with a strip of polythene.

It was found that rotation of the guarded bars caused a change in capacitance, and this was traced to a difference in the diameters and a misalignment of the axes of the bars on either side of the insulating gap. With equal diameters it was still very difficult to assemble the gaps so that there was no variation of capacitance with rotation of the bar. In this case, two measurements with the bar rotated through 180° should eliminate the effect of a radial displacement at the gap. The critical nature of the gap geometry was not appreciated when the model was being constructed, and the imperfect edges of the gap formed a real limitation to its precision. Because of this, no attempt was made to evacuate the capacitor, and corrections for the permittivity of the air were calculated from measurements of temperature, pressure and relative humidity.

The assembly shown in Fig. 5 gives two separate capacitors, bars 1, 2, 4 and 5 forming one and bars 2, 3, 5 and 6 the other. Bars 2 and 5 determine the lengths of both capacitors. The two sections may be connnected in parallel by joining bar 1 to bar 3 and bar 4 to bar 6. In this case the 10 in length gives about 1 pF ($C_0 = 0{\cdot}049\,62$ pF/in). Two separate comparisons with the laboratory's standards of capacitance, which are derived from measurements of resistance and frequency, gave differences of +10 and −3 parts in 10^6 between the calculated and measured capacitances. It is considered that the length uncertainty was about 10 parts in 10^6, and the uncertainty in the electrical measurements was of the same order. The figures indicate that there are no very large systematic errors in the

measurements, but the uncertainty in the final result is too large to justify a more detailed discussion of these determinations. However, the capacitor has been used to provide a great deal of design information for a more precise capacitor. This is of the form shown in Fig. 4. The change in length is measured by a built-in interferometer, and preliminary tests indicate that the uncertainty in the calculated capacitance in electrostatic units should be about 1 part in 10^7.

A large number of electrical measurements are required to relate the value of this precise capacitor (0·25 pF) to a 1-ohm standard resistor. The present limitation to the accuracy of this process is the stability of the capacitors and resistors used for the intercomparison. High-stability low-temperature-coefficient components are being constructed to reduce this limitation. It is possible to obtain a sensitivity of 1 part in 10^7 for all of the electrical measurements, and an overall accuracy of 1 part in 10^6 should be obtained for the value of a 1-ohm resistor in electrostatic units.

(5) CONCLUSION

Tests on an experimental capacitor have confirmed that a cylindrical cross-capacitor may be used as a very precise calculable standard. With reasonable care in construction and assembly, the only uncertainty is that in the defined length. If the length is defined by guard electrodes, the uncertainty may be as much as 1 part in 10^5 for a 10 in length unless very great care is taken in the construction of the insulating gaps. There are a number of techniques for the definition of length which reduce this uncertainty to about 1 part in 10^7.

A practical limitation is the small value of capacitance—1 pF or less—and very few laboratories have equipment for the precise measurement of such small 3-terminal capacitances. However, a relatively simple audio-frequency bridge with a sensitivity of 10^{-7} pF can be constructed for these measurements.[4]

A cylindrical cross-capacitor is very insensitive to films on the surfaces of the electrodes if these films are stable.[5] Because of this it has been suggested that this type of capacitor should be suitable for the accurate measurement of the permittivity of water vapour and other dielectrics which form films on the electrode surfaces.

(6) REFERENCES

(1) THOMPSON, A. M., and LAMPARD, D. G.: 'A New Theorem in Electrostatics and its Application to Calculable Standards of Capacitance', *Nature*, 1956, **177**, p. 888.

(2) LAMPARD, D. G.: 'A New Theorem in Electrostatics with Applications to Calculable Standards of Capacitance', *Proceeding I.E.E.*, Monograph No. 216 M, January, 1957 (**104** C, p. 271).

(3) MCGREGOR, M. C., HERSCH, J. F., CUTKOSKY, R. D., HARRIS, F. J. K., and KOTTER, F. R.: 'New Apparatus at N.B.S. for Absolute Capacitance Measurement' (paper presented at Conference on Electronic Standards and Measurements, Boulder, Colorado, August, 1958).

(4) THOMPSON, A. M.: 'The Precise Measurement of Small Capacitances', *ibid.*

(5) LAMPARD, D. G., and CUTKOSKY, R. D.: 'Some Results on the Cross-Capacitances per Unit Length of Cylindrical 3-Terminal Capacitors with Thin Dielectric Films on Their Electrodes' (not yet published.)

(6) VAN DER PAUW, L. J.: 'A Method of Measuring Specific Resistivity and Hall Effect of Diodes of Arbitrary Shape', *Philips Research Reports*, 1958, **13**, p. 1.

621.317.61:621.385.1

The Institution of Electrical Engineers
Paper No. 2832 E
May 1959
©

A METHOD FOR THE ACCURATE MEASUREMENT OF MUTUAL CONDUCTANCE OF THERMIONIC VALVES

By M. R. CHILD, Associate Member, and D. J. SARGENT, B.Sc.

(*The paper was first received* 11*th April, and in revised form* 13*th November*, 1958.)

SUMMARY

The paper mentions briefly the need for measuring accurately the mutual conductance of thermionic valves at various intervals, and explains in some detail how this measurement can be undertaken. It is difficult to assess the precise magnitude of the errors involved, but, while the inaccuracy in the absolute value of mutual conductance may be as great as 0·25%, it is unlikely that the errors in comparative measurements will exceed 0·1%. The method described is readily adaptable for the measurement of anode conductance and screen-grid amplication factor.

LIST OF SYMBOLS

C = Capacitance.
G = Conductance.
R = Resistance.
V = Signal voltage.
g_a = Anode conductance.
g_m = Mutual conductance.
g_m' = Mutual conductance with screen feedback.
I = Signal current.
r_a = Anode resistance.
r_{g2} = Screen-grid-to-cathode resistance.
α = Ratio (cathode current)/(anode current).
μ = Anode amplification factor.
μ_2 = Screen-grid amplification factor.

(1) INTRODUCTION

Modern techniques and methods of production enable manufacturers to produce thermionic valves which have a high probability of providing faultless service for a number of years. Unfortunately it is not yet possible to ensure that every valve will conform to the required high standard of performance, and it is therefore necessary to submit all potential long-life valves to a test, or series of tests, which will lead to the identification of any substandard valves. The change of mutual conductance with time is considered to be one of the most reliable pointers to the quality of a valve, and the paper is chiefly concerned with describing an accurate method of measuring this parameter. The equipment used in the evaluation of mutual conductance is, however, directly applicable to measurements of anode conductance and of screen-grid amplification factor, and the paper has therefore been extended to include these measurements.

(2) ACCURACY OF MEASUREMENT

The accuracy required in the measurement of a valve parameter such as mutual conductance depends, of course, upon the importance which is to be attached to the result. For example, in a number of applications it is sufficient if the answer is within 5% of the true value; however, when valves which have potential lives of many tens of thousands of hours are to be graded after a life test of, say, 4000 hours, a much higher order of accuracy is required to ensure the necessary degree of confidence in the results. The value of a graph relating g_m with time lies in the interpretation of the slope of the curve between 1000 and 4000 hours, and the decision to accept or reject a valve frequently depends upon this factor. The difference between good and bad valves may, for example, be reflected in slopes of 0·5% and 2·0% per thousand hours. Absolute accuracy of measurement of g_m is not essential provided that any initial errors remain constant for the duration of the life tests, but experience has shown that the maximum error which can be tolerated in the relative measurement of g_m is $\pm 0{\cdot}25\%$.

In the equipment described here the maximum errors in the absolute accuracy do not exceed 0·25%; moreover, since some of the errors involved are substantially constant, the errors between succeeding measurements will not exceed 0·1%.

It will be appreciated from subsequent Sections that greater absolute accuracy can be obtained, if required, by the use of more accurate conductances and by further reducing the resistance of the anode supply.

(3) METHOD OF MEASUREMENT

(3.1) The Valve Test Panel

If a valve is switched off and then on again, as for example when it is transferred from the life-test rack to the measuring equipment, a drift of mutual conductance is observed. To ensure that this drift does not impair the accuracy of the measurement, it is necessary to impose a waiting period of 30 min between the switching-on and the commencement of the measurements. It is quite impracticable, because of the numbers involved, to run each valve individually for this period, and the test panel has therefore been designed to accommodate ten valves (this number was chosen because it represents one unit on the life-test rack). All ten valves are supplied with accurately adjusted d.c. drives, and a switch is provided which selects each valve in turn and connects it into the required measuring circuit without affecting the remaining nine. All the idling valves are supplied with auto-bias from variable potentiometers in their cathode circuits, but the valve under test can, at the option of the operator, acquire its bias in various ways. When the valve is selected for measurement the auto-bias resistor used during the idling period may be left in circuit, or control-grid bias may be substituted. Either of these conditions is suitable for measurement at constant current, but the former is preferred in that it does not disturb the running conditions of the valve, even momentarily; the latter, however, is more suitable when it is necessary to ensure that the anode and screen-grid-to-cathode p.d.'s are not modified by the bias voltage. A third method of biasing which is also available consists in replacing the auto-bias resistor with a calibrated resistor, of 0·1% accuracy, which is variable in steps of 1 ohm from 0 to 1000 ohms; this arrangement is used when measurements are required at fixed cathode-bias resistances.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Child and Mr. Sargent are at the Post Office Research Station.

(3.2) Supplies and Metering

Electronically stabilized power units of the series-regulator type are used to supply the anode, screen-grid and control-grid potentials and to maintain them constant to within 0·1%. The heater supply presented some difficulty initially, but the problem was finally resolved by the development of a stabilized d.c. power unit which provides an output variable between 1 and 5 volts at 8 amp and between 5 and 22 volts at 40 watts. All supply voltages are fed to the valves through a single master switch which ensures that the potentials are applied to the electrodes in the correct sequence.

The current and the voltage parameters are both monitored on high-grade meters which have 6 in scales, knife-edge pointers and anti-parallax mirrors. All meters are provided with shunts, which are variable to the extent of 1% of full-scale deflection and can be used in conjunction with standard meters to ensure a correct calibration. Care has been taken with the layout to ensure that the operator has a non-parallax view of all the meters with only the minimum movement of viewing position. This has been achieved by mounting two of the meters at eye level immediately in front of the operator and by placing the two others on a hinged panel which can be adjusted to provide the correct viewing angle. The applied voltages can readily be adjusted while the meters are being read, and with reasonable care they can be set to within 0·2%; it is apparent from the results obtained, that, in fact, experienced operators keep well within these limits.

(3.3) The Measuring Panel

It will be seen from the following Sections that the same general principles have been adopted for the measurement of mutual conductance, anode conductance and screen amplification factor, and it is therefore possible to utilize the same components for all three measurements. These components are grouped together on the measuring panel, with a switch which arranges them into the correct electrical configuration for any particular measurement and also connects this network, via a screened cable, to the valve test panel.

(4) CIRCUIT FOR THE MEASUREMENT OF MUTUAL CONDUCTANCE

It has long been appreciated*† that the most accurate method of measuring mutual conductance is to apply a small alternating e.m.f. to the control grid of a valve and to compare the resulting current in the anode circuit with a known variable current, so arranged that equality between the two is denoted by zero deflection on a null-balance indicator. The necessity for the precise measurement of small voltages and currents is thus removed, and, if required, provision can be made for a direct reading of mutual conductance in terms of some calibrated control in the balancing-current circuit. In principle, the arrangement adopted conforms to this practice, but it is thought that the particular method used to obtain these results may be of interest.

To ensure correct correlation between measurements of g_m made on the same valve at various times it is essential either to apply exactly the same level of signal to the control grid for every measurement or to make the signal of sufficiently small amplitude that minor variations do not materially affect the value of g_m. For the range of valves so far included in the long-life category, there is no noticeable variation in g_m as a function of the applied grid signal provided that this is always smaller than 0·2 volt d.a.p. In the present arrangement the grid signal has been standardized at 75 mV d.a.p.

* Ballantine, S.: 'The Operational Characteristics of Thermionic Amplifiers', *Proceedings of the Institute of Radio Engineers*, 1919, 7, p. 129.

† Tuttle, W. N.: 'Dynamic Measurement of Electron-Tube Coefficients', *ibid.*, 1933, 21, p. 844.

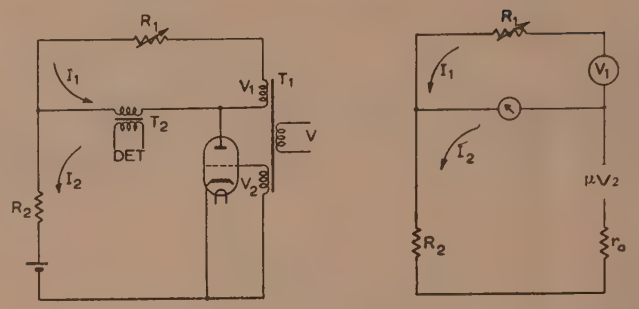


Fig. 1.—Basic circuit for the measurement of mutual conductance, and its equivalent circuit.

The basic circuit for the measurement of g_m is shown in Fig. 1, together with its equivalent network. A 1 kc/s signal voltage, V, is applied to the transformer T_1. One secondary winding on this transformer has a terminal voltage, V_1, which causes a current I_1 to flow via the variable resistor R_1 and the primary of the transformer T_2. The other secondary winding of T_1 applies a voltage V_2 between the control grid and the cathode of the valve under test. The resulting anode current, I_2, which also flows through the primary of T_2, is out of phase with I_1, and therefore the e.m.f. appearing across the secondary of T_2 can be used as a measure of the relative amplitudes of I_1 and I_2.

Considering the case where $I_1 = I_2$, the voltage across T_2 is zero, so that

$$\frac{V_1}{R_1} = \frac{\mu V_2}{r_a + R_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$= \frac{g_m r_a V_2}{r_a + R_2} \text{ since } \mu = g_m r_a$$

i.e.
$$\frac{1}{R_1} = g_m \frac{V_2}{V_1}\left(1 - \frac{R_2}{r_a + R_2}\right)$$

or
$$G_1 = g_m \frac{V_2}{V_1}\left(1 - \frac{R_2}{r_a + R_2}\right) \quad . \quad . \quad . \quad . \quad (2)$$

In the present circuit I_1 is, in fact, varied by the adjustment of four high-stability decade conductances, which are connected in parallel and provide a direct reading of G_1.

The resistance of R_2 in the present circuit is less than 0·5 ohm and r_a is usually greater than 1 kilohm; thus the inaccuracy in the measurement of the absolute value of g_m introduced by neglecting the error term does not exceed 0·05%. Furthermore, it can be seen that this error does not affect the day-to-day measurements, provided that r_a remains sensibly constant. Since this type and magnitude of error is acceptable, eqn. (2) can be rewritten as

$$G_1 = g_m \frac{V_2}{V_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

To enable G_1 to be calibrated in terms of g_m, the ratio V_2/V_1 must be known accurately; the calibration is, of course, simplified if $V_1 = V_2$.

The losses in T_1 make it impossible to maintain a constant relationship between its two secondary voltages for all values of I_1, and the input to the valve has therefore been obtained by means of a low-loss tuned transformer connected between the supply to the balance network and the grid-cathode system of the valve. This arrangement ensures that the amplitudes of the two voltages are equal to within 1 part in 5 000 for all values of I_1.

The improved circuit is shown in Fig. 2, which also includes some other essential refinements and the equivalent network. The extra winding of T_1 is now used in conjunction with R_3 and C_3 to provide an initial phase balance, and it should be noted that the current through these components serves only to sharpen the initial zero balance and does not appear in the final equation.

It was stressed earlier that correctly adjusted parameters are an essential feature of this type of measurement, and therefore an anode-current meter must be included in the circuit. To maintain the accuracy of measurement it is necessary to shunt this meter with a high-grade condenser, and this combination

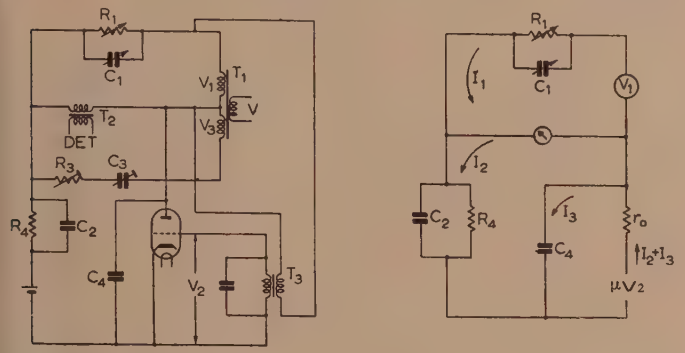


Fig. 2.—Final arrangement for the measurement of mutual conductance, and its equivalent circuit.

appears as R_4C_2 in Fig. 2. For completeness, the anode-to-earth capacitance of the valve has been shown as C_4 and the phase-balance condenser as C_1.

The current flowing through the detector owing to the presence of the signal V_2 at the grid of the valve is given by

$$I_2 = \frac{\mu V_2(1 + j\omega C_2R_4)}{r_a + R_4 + j\omega R_4r_a(C_2 + C_4)} \quad . \quad . \quad . \quad (4)$$

At balance $I_1 = I_2$, and therefore

$$G_1 = g_m\left[1 - \frac{R_4^2 + R_4r_a + \omega^2R_4^2r_a^2C_4(C_2 + C_4)}{(r_a + R_4)^2 + \omega^2R_4^2r_a^2(C_2 + C_4)^2}\right] \quad . \quad (5)$$

The capacitive currents are required solely to provide a sharp null indication and therefore do not appear in this equation.

A typical value for R_4 is 10 ohms, and provided that C_2 is sufficiently large, no serious errors are introduced by the R_4C_2 network. In practice, C_2 has been chosen as 50 μF and the error in g_m, as suggested by this equation, does not exceed 0·1%, provided that r_a exceeds 1 kilohm. Furthermore, this error is important only when the absolute value of mutual conductance is required, and it does not affect the relative accuracy of measurements made on the same valve at different times.

The circuit has been treated in terms of a triode, for simplicity, but it is equally applicable to a pentode, provided that the resistance of the screen-grid supply, R_{g2}, is kept small. The effect of resistance here is given by

$$g_m' = \frac{g_m}{1 + \frac{g_m}{\mu_2}(\alpha - 1)R_{g2}} \quad . \quad . \quad . \quad . \quad (6)$$

The resistance of the screen supply has been kept below 1 ohm in the present equipment and the errors from this source can therefore be neglected.

(5) MEASUREMENT OF ANODE CONDUCTANCE

By a simple rearrangement the components used for measuring g_m are connected to form a conductance-measuring bridge in which the anode conductance of the valve forms the unknown arm. To enable the decade conductance, already chosen for the measurement of g_m, to cover the range of anode conductances normally encountered, it is necessary to provide the bridge with two ranges. The method adopted to obtain these is shown in

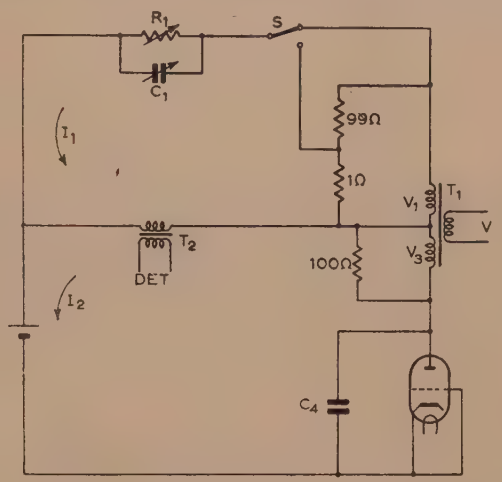


Fig. 3.—Basic circuit for the measurement of anode conductance.

Fig. 3, where the basic circuit used for this measurement is outlined.

The accuracy of this arrangement depends to some extent upon the value of g_a, and over the range 0·5–10 000 micromhos the errors do not exceed 1%.

(6) MEASUREMENT OF SCREEN-GRID AMPLICATION FACTOR

A method for measuring the screen-grid amplification factor, μ_2, of a pentode is depicted basically in Fig. 4. The same com-

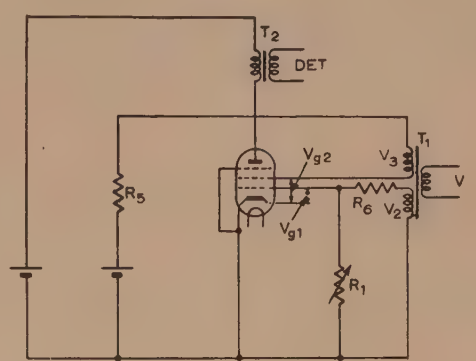


Fig. 4.—Basic circuit for the measurement of screen-grid amplification factor.

ponents have again been used, and in this instance the two secondary windings of T_1 apply out-of-phase voltages V_2 and V_3 to the control grid and the screen grid, respectively, of the valve. The resulting anode current which flows through the primary of T_2 is dependent upon the relative magnitudes of V_2 and V_3 and upon the value of μ_2.

In the circuit shown in Fig. 4 it is seen that the screen-grid-to-cathode voltage, V_{g2}, is given by

$$V_{g2} = \frac{V_3r_{g2}}{R_5 + r_{g2}} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where r_{g2} is the screen-grid-to-cathode resistance of the valve.

Similarly, the control-grid-to-cathode voltage, V_{g1}, is given by

$$V_{g1} = \frac{V_2 R_1}{R_1 + R_6} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

If under these conditions the resultant anode current is zero, then, by definition,

$$\mu_2 = \frac{V_{g2}}{V_{g1}}$$

$$= \frac{V_3}{V_2}\left(1 - \frac{R_5}{R_5 + r_{g2}}\right)\left(1 + \frac{R_6}{R_1}\right) \quad . \quad . \quad . \quad (9)$$

The omission of all inter-electrode capacitances is fully justifiable at the operating frequency of 1 kc/s. R_5 is about 0·2 ohm, and for most valves the term in the first bracket approximates closely to unity. The design of the transformer T_1 is such that $V_3/V_2 = 1$ to within 0·2% for all likely values of R_1, and eqn. (9) can therefore with reasonable accuracy be reduced to

$$\mu_2 = 1 + R_6 G_1 \quad . \quad . \quad . \quad . \quad . \quad (10)$$

since $G_1 = 1/R_1$.

The choice of R_6 is dependent upon the range of μ_2 which has to be covered, and it determines the multiplication factor for the dial readings of G_1; therefore, in effect, $\mu_2 = 1 + G_1$, where G_1 is in suitable units.

The slight inconvenience of having to add a constant to the measured value of μ_2 is thought to be more than compensated by the simplicity of the arrangement, by the use of only one set of measuring dials for g_m, g_a and μ_2 and by the accuracy of measurement.

(7) AUXILIARY EQUIPMENT

In addition to the equipment already described the measuring set embodies a stable 1 kc/s oscillator and a null-balance detector. The output from the Van der Pol type of oscillator is variable from 0–10 volts d.a.p. for experimental purposes, but for routine work one of three preset levels is automatically chosen when the circuit-selecting switch is adjusted to either g_m, g_a or μ_2.

The detector is a narrow-band 3-stage amplifier, of low noise antimicrophony construction, in which the output is fed via a crystal rectifier to an ammeter with a full-scale deflection of 100 μA. A limiting device has been incorporated to prevent the meter deflection exceeding full scale. The sensitivity of the amplifier is such that a 2 μA deflection on the meter is obtained for an input current of 0·01 μA.

The whole equipment is flush-mounted on two racks, a writing ledge is provided at a convenient height, and the meters and controls are arranged to ensure a comfortable operating position. The racks are enclosed on the sides and back in a cabinet type of construction which provides easy access to all parts.

(8) EXPERIENCE

Various forms of the measuring set described in the paper have been in almost daily use at the Post Office Research Station for some 10 years. The instruments have fully justified the confidence placed in them, and have been invaluable in the assessment of the quality of thermionic valves which were being life-tested prior to their inclusion in such schemes as the transatlantic submerged-repeater telephone cable.

At least one other manufacturer concerned with the production of long-life valves is using this form of measuring set.

It is now thought, however, that a worth-while saving in operating time could be introduced without any loss of accuracy by adapting the present measuring set and by modifying the valve life-test panels so that valves could be measured without being first removed from the life-test racks. The possibility of developing such a scheme is receiving active consideration at the Post Office Research Station.

(9) ACKNOWLEDGMENTS

The authors wish to express their thanks to All Power Transformers, Ltd., for technical assistance in the design of the special power units and also for their assistance in engineering the equipment on a commercial basis. Acknowledgment is also made to the Engineer-in-Chief of the Post Office for permission to make use of the information contained in the paper.

621.396.677

The Institution of Electrical Engineers
Paper No. 2807 E
May 1959

THE PERFORMANCE OF DIRECTIVE AERIALS IN COMPLEX U.H.F. FIELDS

By J. A. SAXTON, D.Sc., Ph.D., Member, and B. N. HARDEN, M.Sc.

(*The paper was first received 9th September, and in revised form 25th October,* 1958.)

SUMMARY

The paper describes an investigation, at frequencies of 580 and 904 Mc/s, of the apparent gains relative to a half-wavelength dipole of a number of directive aerials at a typical selection of receiving sites in urban and rural areas. It is shown that variations of the apparent gains occur, as a result of the complexity of the radio field patterns generally found at such sites, to an extent which may be significant in the estimation of the coverage of transmitters operating in the broadcasting bands IV and V.

(1) INTRODUCTION

In assessing the usefulness of ultra-high frequencies for the provision of a broadcast service such as television, a number of factors depending upon propagation characteristics at these frequencies must be taken into account. For example, it is essential to know the coverage produced by the ground wave over terrain of various kinds, and in addition, if the use of the same frequency by two or more transmitters is envisaged, the interfering effects due to long-distance tropospheric propagation must also be understood. It is with one aspect of the first of these two problems that the paper is concerned.

Now, although it is possible to obtain a good statistical evaluation of the coverage produced by a u.h.f. transmitter in a given area,[1] the precise nature of the field pattern at any specified receiving site cannot be predicted, and often the local field variations are quite complex. This complexity is mainly associated, in the service area, with the effects of diffraction and re-radiation produced by objects such as hills, buildings, trees, and electricity supply and telephone wires in the neighbourhood of the site. Since, for a worth-while service area to be achieved, it is much more necessary to use directive aerials at ultra-high frequencies than is generally the case at very high frequencies, if prohibitively large transmitter powers are not to be required,[2,3] it is important to realize that these aerials may have a performance in typically occurring complex fields which differs appreciably from that which they would have for an incident plane-wave field. Whilst one of the desirable features of a directive aerial may be its ability to exclude delayed signals, and indeed this may be its most important attribute well inside the service area, the realization of adequate gain is of paramount importance towards the edge of the service area. In view of the serious consideration being given at the present time to the possibility of using frequencies in the broadcasting band V (610–960 Mc/s) for television, investigations have been carried out at the Radio Research Station to see how far the performance of typical directive receiving aerials is influenced by field complexity at ultra-high frequencies; the feature which has in particular been examined is the effective gain of these aerials relative to a half-wavelength dipole.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is an official communication from the Radio Research Station, Department of Scientific and Industrial Research.

(2) EXPERIMENTAL PROCEDURE

The gain measurements were made at frequencies of 580 and 904 Mc/s on horizontally polarized signals radiated from transmitters at the Radio Research Station, Slough, and the receiving sites examined, at distances up to 20 km, embraced a wide variety of urban and rural conditions. Three types of directive aerial were used: Yagi arrays, and a half-wavelength dipole associated with either a plane reflector or a corner reflector, the reflectors being of aluminium sheet. Field-strength measurements were also made with a simple half-wavelength dipole for reference purposes.

Arrangements were made to attach the various receiving aerials in turn, at a height of 9 m above ground, to a rotatable wooden mast carried on a motor van. A balance-to-unbalance transformer was fitted to each aerial, and the same coaxial cable was used in all cases to connect the aerials to a calibrated receiver. Comparisons of the received signals were made by adjustments of a piston attenuator in the intermediate-frequency section of the receiver to maintain a constant output. Observations of the effective aerial gain were made at each site under two sets of conditions:

(*a*) With each aerial oriented in the direction for which the maximum signal would have been expected, having regard to the true bearing of the transmitter from the receiver.
(*b*) With each aerial set to give the actual maximum signal.

The gains of the aerials were also measured on a clear open site, using a local transmitter, where experience had shown that it was possible to establish a substantially plane-wave field at the receiver. Horizontally polarized radiation was used, and measurements were made with the various aerials mounted on the mast in the same manner as for the remainder of the investigation.

It is considered that, under the best receiving conditions with steady fields, gain measurements could be made with an error not exceeding $\pm 0{\cdot}5$ dB: on sites where the field was somewhat variable, owing, for example, to the movement of trees in the vicinity, the error was probably liable to be greater, though it is difficult to give an accurate estimate of what it might have been.

(3) RESULTS

The directive aerials investigated at the two frequencies are listed in Table 1, with their measured plane-wave gains relative to a half-wavelength dipole. It should be noted that no attempt was made to optimize the design of any of these aerials, for example, to obtain the maximum possible gain for a Yagi array having a given number of elements, since all that was required was a number of aerials covering a reasonably wide range of gain, typical of that likely to obtain in practical conditions of domestic reception. It is, however, unlikely that a significantly superior performance could in fact have been achieved with any given type of aerial, except possibly with those based on the corner reflector. All of the aerials were adjusted to have an impedance of 70 ohms, which was also the characteristic impedance of the coaxial feeder.

Table 1

AERIALS INVESTIGATED

	Frequency	Type of aerial	Plane-wave gain
	Mc/s		dB
1	580	Plane reflector, 1 wavelength (λ) square. $\lambda/2$ dipole, $0{\cdot}2\lambda$ in front of reflector	6·0
2	580	Corner reflector A, 90° angle, sides $0{\cdot}7\lambda \times 1\lambda$. $\lambda/2$ dipole, $0{\cdot}35\lambda$ from apex*	7·5
3	580	Yagi array A. Reflector, folded dipole, 3 directors $0{\cdot}2\lambda$ spacing	7·75
4	580	Yagi array B. Reflector, folded dipole, 8 directors $0{\cdot}18\lambda$ spacing	9·0
5	580	Yagi array C. Reflector, folded dipole, 8 directors $0{\cdot}34\lambda$ spacing	10·75
6	904	Plane reflector, 1λ square. $\lambda/2$ dipole, $0{\cdot}2\lambda$ in front of reflector	6·0
7	904	Yagi array D. Reflector, folded dipole, 3 directors $0{\cdot}2\lambda$ spacing	7·5
8	904	Corner reflector B, 90° angle, sides $0{\cdot}7\lambda \times 1\lambda$. $\lambda/2$ dipole, $0{\cdot}35\lambda$ from apex*	7·75
9	904	Yagi array E. Reflector, folded dipole, 8 directors $0{\cdot}18\lambda$ spacing	8·5
10	904	Corner reflector C, 90° angle, sides $1{\cdot}1\lambda \times 1{\cdot}5\lambda$. $\lambda/2$ dipole, $0{\cdot}35\lambda$ from apex*	9·75
11	904	Yagi array F. Reflector, folded dipole, 13 directors $0{\cdot}34\lambda$ spacing	13·0

* The longer dimension is measured along the apex, which is horizontal.

Fig. 1 shows the statistical variations of the gains of aerials 1–5 at a frequency of 580 Mc/s, when oriented for maximum received signal at each site; Fig. 2 gives similar distributions for aerials 6–11 at 904 Mc/s. The receiving sites used were not

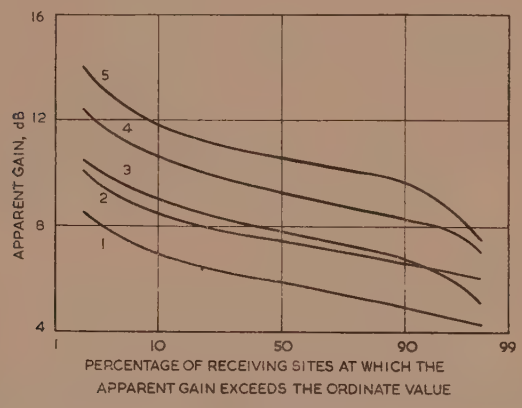


Fig. 1.—Variation of apparent gain of aerials 1–5 at a frequency of 580 Mc/s.

the same for the two frequencies, nor for all of the aerials at each frequency, but a large and representative selection of sites was examined in each case. It will be noted from Figs. 1 and 2 that the form of the distribution curve is much the same for all of the aerials investigated. Values above and below the plane-wave gains were observed, but for each aerial the median value of the distribution was within $\pm 0{\cdot}5$ dB of the appropriate plane-wave gain. However, although the overall distributions were similar, the relative gains of the aerials, one to another, varied widely at individual sites.

The difference in gains relative to that of a dipole at the same location, measured by directing the aerial for maximum signal and in the direction determined by the true bearing of the transmitter, is illustrated by Fig. 3, which presents results obtained

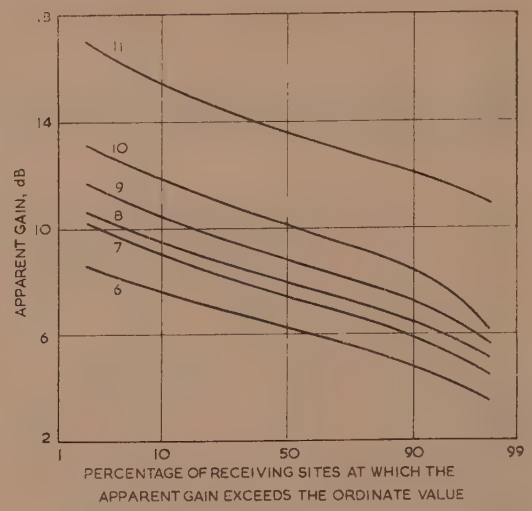


Fig. 2.—Variation of apparent gain of aerials 6–11 at a frequency of 904 Mc/s.

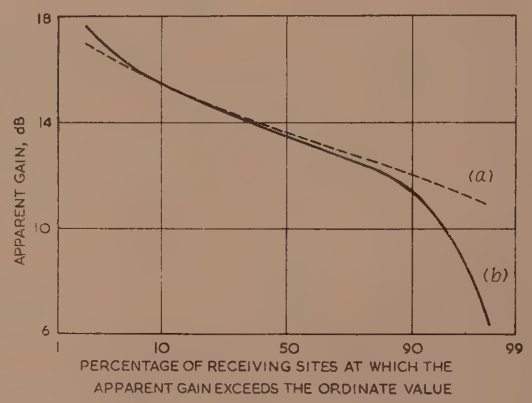


Fig. 3.—Variation of apparent gain of Yagi array F with orientation and location (frequency 904 Mc/s).
(*a*) Aerial oriented for maximum signal.
(*b*) Aerial oriented for true bearing of transmitter.

at 904 Mc/s using aerial No. 11 (Yagi array F). This was the most directive of all the aerials, and the difference between the statistical distributions obtained for the two orientations was greater than in any other case: the direction for the maximum received signal was at times as much as 20° away from the true bearing of the transmitter.

Detailed measurements were made at two sites to investigate the manner in which the apparent gain varied with small displacements of the aerial, Yagi array F at 904 Mc/s again being used, and the results are shown in Figs. 4(*a*) and (*b*). The first site, to which Fig. 4(*a*) refers, was on a road bordered by two-storeyed houses with trees and overhead wires in the vicinity; the second site [Fig. 4(*b*)] was in a nearby district amongst bungalows, but there were no trees or wires and the aspect was open. Observations were made for both of the orientations of the aerial defined above, and it will be seen that the variations were much greater on the obstructed site. Similar results were obtained at other locations, and they may thus be regarded as typical of what will occur in practice, though differences in performance between a directive aerial and a dipole are obviously likely to be relatively less pronounced for aerials of lower directivity.

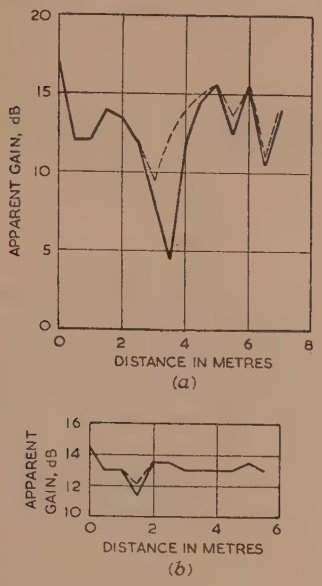


Fig. 4.—Variation of the apparent gain relative to a half-wavelength dipole of Yagi array F (904 Mc/s) over short distances.

(*a*) On an obstructed site.
(*b*) On an unobstructed site.

---- Aerial oriented for maximum signal.
—— Aerial oriented for true bearing of transmitter.
Measurements made at intervals of 0·5 m.

(4) DISCUSSION

The effect of local obstacles on the apparent gain of a directive aerial is clearly shown in Fig. 4(*a*), and the complex field at any given receiving site is presumably the resultant of a number of components due to diffraction and re-radiation by neighbouring obstacles, in addition to the directly transmitted component. In such circumstances a horizontal half-wavelength dipole receiving horizontally polarized waves may be affected equally by components arriving not only from the front, but also from the rear and from above and below. On the other hand, all of the directive aerials examined included a reflector and were predominantly influenced by radiation coming from the forward direction, though in many cases the main beam of the radiation pattern was quite broad. Considering the simple case of two field components, one arriving from the forward direction and one from the rear, if they are out of phase at the receiving point the signal obtained using only a dipole will be relatively low: with a directive aerial, however, some appreciable discrimination against the component from the rear will occur, and the apparent gain of this aerial relative to the dipole will be greater than the plane-wave value. If the same two components are in phase at the receiving point, the apparent gain will be less than this value. With the increasing field complexity caused by a larger number of components of varying amplitude and phase, it is not surprising that there is a considerable spread in the observed apparent gain relative to a dipole, both above and below the plane-wave gain. Furthermore, two directive aerials having substantially different radiation patterns and plane-wave gains may well not give the same difference in performance in a complex field, and their relative performance may vary from site to site depending upon the number and directions of arrival of the various signal components at these sites.

The curves of Figs. 1 and 2 show that the apparent gains of the aerials examined, expressed in decibels, all follow approximately a normal distribution between the values exceeded at 10% and 90% of the receiving sites, with the median lying close to the appropriate plane-wave gain; the standard deviations being 0·75 dB and 1·25 dB at 580 and 904 Mc/s respectively. There are a small number of sites where the effective gain of the directive aerial relative to a dipole is very small when the aerial is oriented according to the true bearing of the transmitter, and, at some of these at least, considerable improvement could be obtained by orienting the aerial for maximum signal. (The statistical character of Fig. 3 must be emphasized; at some sites, even though the apparent gain is relatively low, the directions for maximum signal and the bearing of the transmitter may coincide.) The greatest observed improvement thus achieved was 8 dB with Yagi array F at 904 Mc/s.

(5) CONCLUSIONS

Measurements of the apparent gain relative to a half-wavelength dipole of a selection of directive aerials, typical of those likely to be used in practice for domestic reception in the television broadcasting bands IV and V, have been made at a wide variety of receiving sites. The observations show that the performance of these aerials may vary considerably, and that, whilst the median value of the gain of a given aerial is close to that which would be expected under the ideal conditions of plane-wave reception, at a significant fraction of sites the gain of the aerial is markedly less than the plane-wave value, whether the aerial is oriented according to the true bearing of the transmitter or for maximum reception. This variation in aerial performance is a factor which should be taken into account in the statistical evalution of the service area of a u.h.f. broadcasting transmitter.

(6) ACKNOWLEDGMENT

The work described in the paper was carried out as part of the programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

(7) REFERENCES

(1) SAXTON, J. A.: 'Basic Ground-Wave Propagation Characteristics in the Frequency Band 50–800 Mc/s', *Proceedings I.E.E.*, Paper No. 1601 R, January, 1954 (**101**, Part III, p. 211).

(2) SAXTON, J. A.: 'Television Coverage', *Wireless World*, 1954, **60**, p. 173.

(3) SMITH-ROSE, R. L., and SAXTON, J. A.: 'U.H.F. Television Broadcasting', *ibid.*, 1955, **61**, p. 343.

621.396.67

The Institution of Electrical Engineers
Paper No. 2849 E
May 1959
©

HEMI-ISOTROPIC RADIATORS FOR THE S- OR X-BAND

By E. G. A. GOODALL, M.Sc.

(*The paper was first received 28th June, and in revised form 1st December*, 1958.)

SUMMARY

An aerial having approximate hemi-isotropic properties has been constructed on the principle that a dielectric rod will act as a guiding medium for electromagnetic energy. Using this principle, a broadband, shaped dielectric element has been developed, which, when placed at the aperture of an open-ended circular waveguide, radiates with hemi-isotropic cover over a 20% frequency band.

(1) INTRODUCTION

In aerial theory the isotropic source is regarded as an idealized concept against which all other aerials are compared. An aerial possessing hemi-isotropic properties would undoubtedly be valuable in certain practical applications such as beacons and aircraft antennae. The paper describes a method of obtaining such an aerial. In practical terms, a hemi-isotropic aerial may be described as a radiating element transmitting circularly polarized waves equally in all directions over a hemisphere. Some methods of obtaining hemi-isotropic cover have been suggested. One is to use a single element which radiates the required patterns, whilst a second involves the principle of gap filling. In such an arrangement the radiation from a broadside-fire aerial, whose radiation pattern has a doughnut shape, is added to the radiation from an end-fire aerial whose radiation pattern is a beam. The gap in the pattern of one aerial is filled by the radiation from another. Assuming the physical difficulties of such an arrangement have been overcome, the correct phasing of the two aerials presents a major problem which is not easy to solve. Knudsen[1] has proposed using a rearranged form of turnstile array, and this appears to be the most promising of the methods involving phasing.

In the paper, a broadband matched single element is described, whose radiation pattern approaches that of an idealized hemi-isotropic aerial. The element which has these characteristics has an approximate mushroom shape, and was manufactured from Perspex, the machining properties of which made it an obvious choice. The mechanism of radiation from an element of this particular shape is not easily determined and thus the design has been entirely empirical.

(2) DESIGN PRINCIPLES

In the present design, circular polarization was obtained by propagating an H_{11} mode along a circular waveguide and through a dielectric circularizer. When an open-ended waveguide is propagating circularly polarized waves it can be expected that circularly polarized waves will be received when a dielectric element is present, at all events along the axis.

It is known that a dielectric rod will act as a guiding element;[2] if such a rod is placed with one end in the mouth of a vertical waveguide and if its other end is bent towards the horizontal, it should draw energy from the vertical direction towards the horizontal. This idea suggested that the whole system could be made symmetrical by using the solid of revolution generated when the element is rotated through 360° about the vertical waveguide axis.

According to the theory given by Brown and Spector[3] the above outline of the radiation process from a dielectric rod is not complete. To the radiation emitted from the free end of the dielectric rod must be added the radiation from the open end of the waveguide. The percentage power radiated from the open end of a rectangular waveguide from which a rectangular dielectric rod is protruding has been calculated by Clarke,[4] who used a transmission-line theory approach to obtain his results. It may not be entirely correct to state that the radiation pattern of the element, which is described later, can be accounted for by these two theories, which have only been developed for a straight rectangular dielectric rod. However, a bending of the rod is not expected to cause a major effect, and thus these theories may indeed be the basis of an explanation of the observed radiation patterns. It appears, therefore, that to obtain an element with a hemi-isotropic radiation pattern, its dimensions must be such that:

(*a*) The correct phasing between the two forms of radiation pattern exists.

(*b*) The correct percentage of power is radiated in each pattern.

(3) RADIATORS

(3.1) X-Band Radiator

Using the above principles, a shaped Perspex aerial was designed and manufactured to operate in the X-band.[5] A vertical section of the aerial system is shown in Fig. 1. In order to introduce as small a discontinuity as possible, the dielectric element was gradually tapered back into the circular waveguide. Between the tapered section and the actual radiating element there was a small parallel section of dielectric, which completely filled the guide and acted as a support to the aerial.

The aerial was inserted in the mouth of a $\frac{7}{8}$ in internal diameter circular waveguide and was tested on a short free-space site, where a small pyramidal horn, placed at a distance of 6 ft from the aerial under test, acted as the transmitter.

In the measurement of the radiation pattern from the dielectric aerial, circular polarization was replaced by linear polarization. With the waveguide axis horizontal and the aerial rotating in a horizontal plane, polarization perpendicular to this plane was transmitted and the latitudinal E_ϕ component was measured. Changing the polarization to horizontal and still rotating the aerial in a horizontal plane, the longitudinal E_θ component was then measured. The definition of the E_θ and E_ϕ components may be seen more clearly in Fig. 2.

Altogether seven models of the same fundamental shape but with varying parameters were tested over the frequency range 8·8–10 Gc/s. A variation of the dimensions of the dielectric aerial produced the following effects:

(*a*) Increasing the overall diameter produced a radiation pattern composed of a large axial beam with small side beams.

(*b*) A variation of the flare angle from 30° to 45° produced a pattern composed of three beams, but little difference resulted in changing the angle from 30° to 20°.

(*c*) The length AB was eventually found to be critical and at X-band frequencies a value of 0·7 in alone gave good results.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.

Mr. Goodall is at the Research Laboratories of Marconi's Wireless Telegraph Co., Ltd.

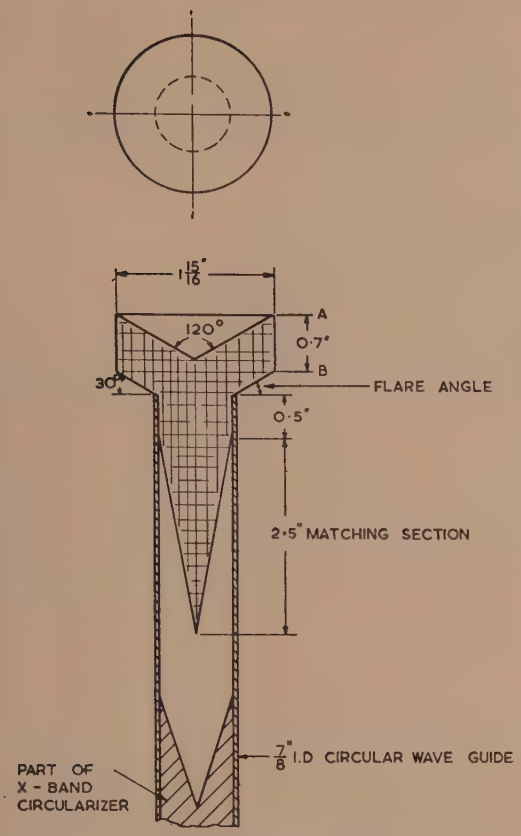


Fig. 1.—X-band omni-directional dielectric aerial.

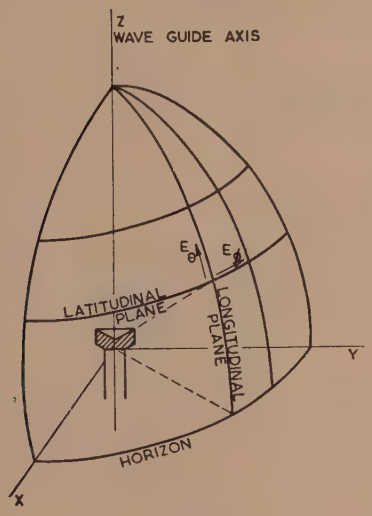


Fig. 2.—Radiation pattern co-ordinate system.

Fig. 3 shows the pattern in the vertical plane for the X-band model at $\lambda = 3{\cdot}2$ cm. The azimuth patterns for the E_θ and E_ϕ components were measured and both were found to be omni-directional with the E_θ component approximately 6 dB down on the E_ϕ component. The final X-band version is shown in Fig. 1. Over the band of frequencies 8·8–10 Gc/s the radiation pattern of this aerial did not change appreciably from that shown in Fig. 3.

The gain of the aerial for the E_ϕ component in a direction of 40° from the waveguide axis was measured against a standard X-band gain horn and was approximately 5 dB above an isotropic

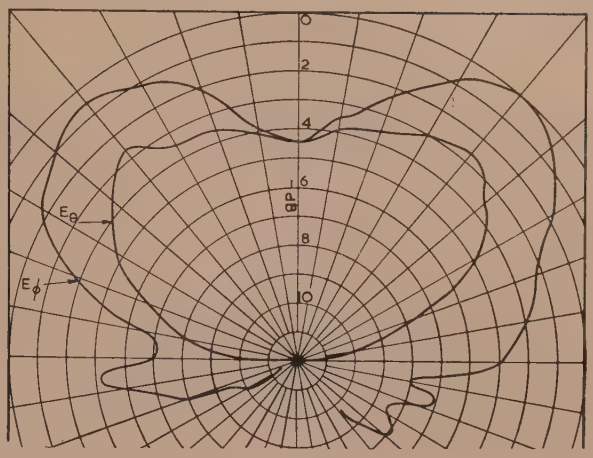


Fig. 3.—Elevation patterns of dielectric aerial. $f = 9{\cdot}375$ Gc/s.

source. The gain horn used was a pyramidal horn of aperture 18·9 × 14·9 cm and 27·5 cm in length, the measured gain being 22 dB above an isotropic source.

The v.s.w.r. of the aerial assembly, which included a rectangular-to-circular waveguide transition but no circularizer, was measured as 0·88 at a wavelength of 3·2 cm. The crystal used in these measurements had previously been calibrated against a standard variable attenuator.

The effect of a ground plane which would be present in any practical application was investigated, for it was thought that its presence might cause a serious deterioration of the observed pattern. With a plane of approximately 10 wavelengths in diameter placed directly behind the aerial, a ripple of −5 dB in depth was introduced into the E_θ component (Fig. 4). The E_ϕ

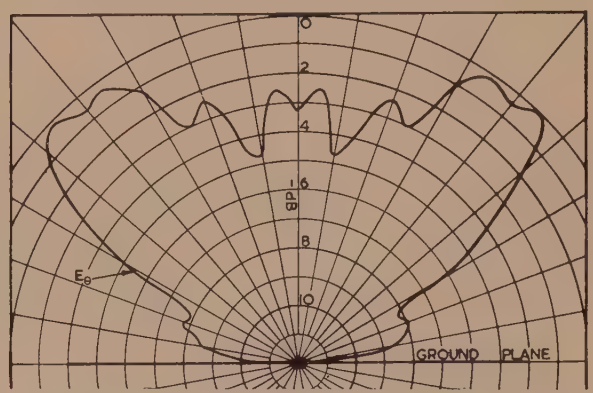


Fig. 4.—E_θ pattern of X-band model with ground plane at mouth of waveguide. $f = 9{\cdot}375$ Gc/s.

component was not measured, but owing to the similarity between the two components it would be expected that the effect would be the same.

(3.2) S-Band Radiator

As the investigation of the X-band aerial was limited in the range of frequencies over which it could be tested, it was decided to build an S-band model to determine the frequency band over which this particular type of radiating element could be used satisfactorily. The dimensions of the S-band model were determined by using the X-band aerial dimensions and scaling by a factor of 2·8, this figure being obtained from the ratio of the two free-space wavelengths at which the aerials were to operate.

As in the X-band model, the critical length AB was modified

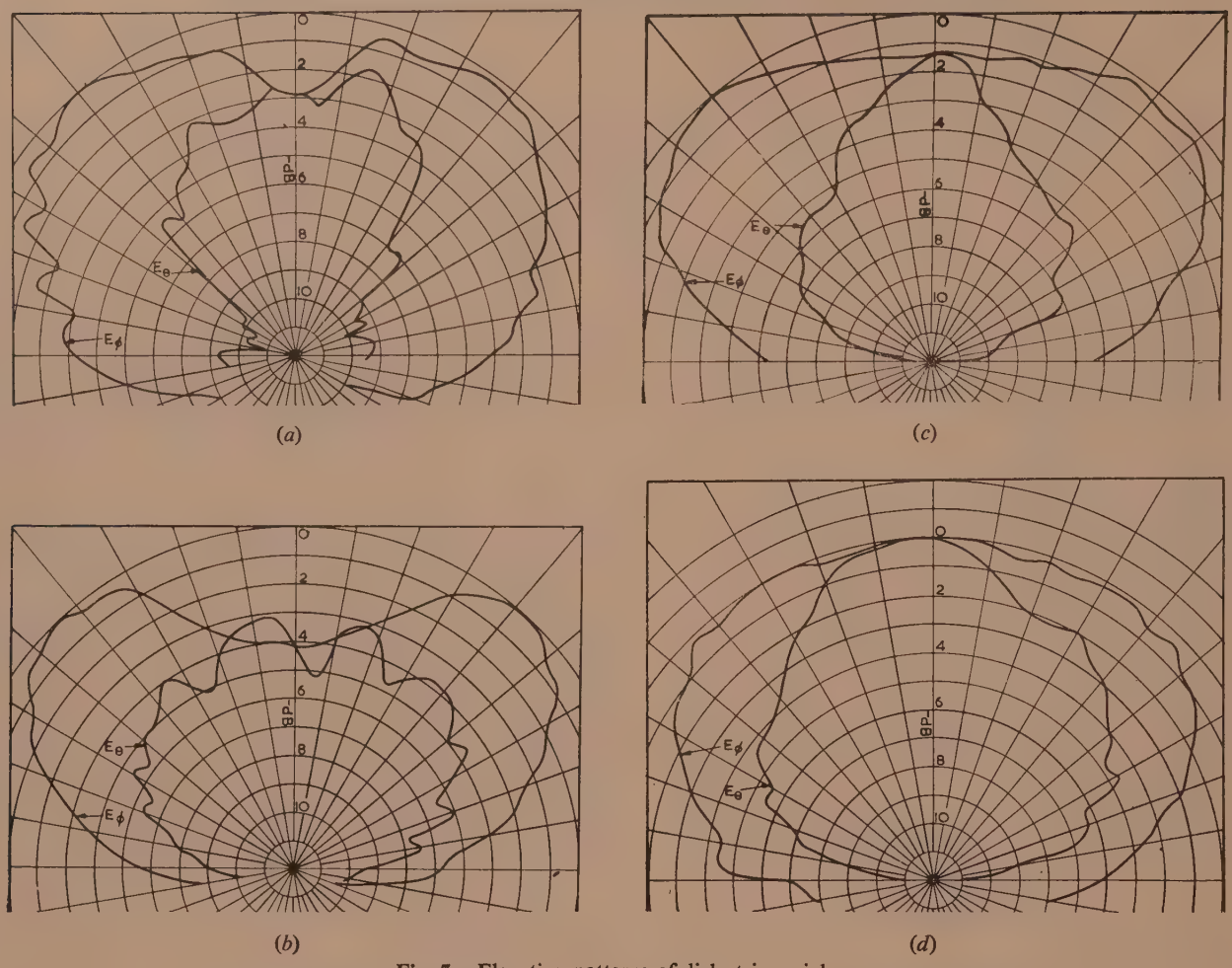


Fig. 5.—Elevation patterns of dielectric aerial.

(*a*) $f = 3{\cdot}33$ Gc/s.
(*b*) $f = 3$ Gc/s.
(*c*) $f = 2{\cdot}86$ Gc/s.
(*d*) $f = 2{\cdot}75$ Gc/s.

during the experimental investigations. A variation from $1\frac{1}{4}$ in to 2 in was made, but there was no marked effect such as occurred at X-band frequencies. A length of $1\frac{7}{8}$ in was finally taken, which gave the most satisfactory radiation pattern. A calibration of the receiver crystal in conjunction with a d.c. amplifier was made at every 0·5 cm change in wavelength from 8 to 11·0 cm. This was necessary because of the considerable change in crystal plus crystal-holder response with frequency.

The aerial was tested over a 25 ft field site using a pyramidal horn aperture, $5\frac{1}{2}\lambda \times 7\frac{1}{2}\lambda$, as a receiver. The aerial under test, acting as the transmitter, was placed with its waveguide axis horizontal and was rotated in a horizontal plane. The oscillator gave a frequency coverage from 2·75 to 3·75 Gc/s. The E_θ and E_ϕ component elevation patterns were measured over a frequency range of 2·75–3·3 Gc/s and are shown in Fig. 5 on a decibel scale.

Although plotting the polar diagrams in field strength would have given a clearer indication of the energy below the horizon, there is no justification for doing so since the energy there is very small and is of no interest if the aerial is used as a beacon.

A measurement of the match of the complete experimental arrangement was made over the frequency range, and the v.s.w.r. was found to be not worse than 0·6. From Fig. 5(*d*) it may be observed that a further extension of the working limit of the aerial is certainly possible, for at this frequency of 2·75 Gc/s the best polar diagram is obtained.

Polarization characteristics of the radiated energy of these dielectric models are as follows:

Circular polarization appears at the zenith, with elliptical polarization between the zenith and the horizon.

On the horizon the horizontally polarized component is 6 dB above the vertically polarized component.

(4) ACKNOWLEDGMENTS

The author wishes to thank the Engineer-in-Chief of Marconi's Wireless Telegraph Co., Ltd., for permission to publish the paper.

(5) REFERENCES

(1) KNUDSEN, H. L.: 'The Quadruped Antenna', *Acta Polytechnica*, 1956, No. 200 (Electrical Engineering Series, **7**, No. 4).

(2) SCHELKUNOFF, S. A.: 'Hyper-Frequency Wave Guides', *Bell System Technical Journal*, 1936, **15**, p. 310.

(3) BROWN, J., and SPECTOR, J. O.: 'The Radiating Properties of End-Fire Aerials', *Proceedings I.E.E.*, Paper No. 2216 R, January, 1957 (**104** B, p. 27).

(4) CLARKE, R. H.: 'A Method of Estimating the Power Radiated directly at the Feed of a Dielectric-Rod Aerial', *ibid.*, Paper No. 2395 R, September, 1957 (**104** B, p. 511).

(5) British Patent No. 808941.

621.372.8:621.317.34

The Institution of Electrical Engineers
Paper No. 2852 E
May 1959
©

THE DESIGN AND TESTING OF INTEGRALLY CONSTRUCTED WAVEGUIDE ASSEMBLIES

By G. CRAVEN and V. H. KNIGHT.

(*The paper was first received 25th August, and in revised form 8th December*, 1958.)

SUMMARY

Problems associated with integrally constructed waveguide assemblies are discussed, the advantages of this form of construction are shown, and suitable components and methods of construction are described. Complex assemblies of this type present special testing problems which are solved by employing a plug-in reflectometer, together with a swept-frequency reflection display. These principles are applied to a typical microwave link repeater, and the alignment of the signal-input section is described in detail. An overall performance with less than 2% reflection is confirmed by independent measurement. A swept-frequency display receiver of high sensitivity is needed for display; the design is briefly described and used to illustrate another example of integral construction.

LIST OF PRINCIPAL SYMBOLS

A_1 = Amplitude coefficient of wave excited in the same direction as the wave in the main guide.
B_1 = Amplitude coefficient of wave excited in the opposite direction to the wave in the main guide.
f = Frequency, Mc/s.
δf = Incremental frequency displacement.
E = Electric field.
H = Magnetic field.
P = Radio-frequency power.
m = Principal magnetic polarizability of the hole.
p = Principal electric polarizability of the hole.
l_c = Slot length.
l = Separation of post centres.
a = Guide width.
b = Guide height.
d = Coupling-hole diameter.
r = Post radius.
t = Guide-wall thickness.
x = Co-ordinate along broad dimension.
z = Longitudinal co-ordinate.
$\alpha = x/a$.
γ = Propagation coefficient.
λ = Free-space wavelength, cm.
λ_c = Cut-off wavelength, cm.
λ_g = Guide wavelength, cm.

(1) INTRODUCTION

Considerable effort has been devoted in recent years to reducing the size and high cost of microwave assemblies, and this has led to the introduction of microstrip in a number of fields where it has been highly successful.[1, 2] Construction in this medium is inherently less bulky, and lends itself to mass production by printed-circuit techniques. However, the medium has certain limitations, it being unsuitable for high-power applications or where high isolation between components is required. Certain components, such as high-Q resonant cavities and narrow-band filters, are not easily realized, and, in general, a low-reflection performance is not readily obtained. Thus waveguide construction is frequently essential, and it is desirable that the waveguide should be as compact and easy to manufacture as possible. This can be achieved by methods that have some parallel with microstrip.[3, 4] One of the advantages of microstrip is its planar form which facilitates construction and installation. Similar benefits are obtained if the waveguide routes are reduced from the customary three planes to only two: thus components of a convenient shape are more easily constructed. If, in addition, the complete unit is designed so that all the components are combined in one integral assembly, much of the connecting guide and weighty flanges can also be eliminated. This departure from standard practice permits radically different methods of fabricating the assembly. It also has the advantage, where only small reflections can be tolerated, of eliminating reflection at a number of flange joints.

Integral construction, in the form discussed, presents three separate problems: suitable components must be designed, a method of construction devised, and a technique developed for the *in situ* alignment and testing of components that are electrically inaccessible, i.e. components that are separated from an entry port by at least one other component. These problems, which are to be discussed, are best illustrated by examples, the main choice being a microwave-link repeater. A typical microwave-link repeater employs a number of filters, several mixers and probably some directive couplers, and the complete assembly is a complex one. The tuning of filters is critical, and the signal circuits must be matched for very low reflection if satisfactory operation is to be obtained. Integral construction results in a large number of components being electrically inaccessible. The problem is thus a formidable one, and a technique that is satisfactory in this case is likely to be suitable for a great many applications.

(2) MICROWAVE-REPEATER DESIGN

The type of repeater to which the integral-construction technique is applied is illustrated in the block schematic of Fig. 1, and will be briefly described. The chief feature of the scheme is that both receiver and transmitter employ a common microwave oscillator. The local oscillator of the receiver is derived from the common oscillator by frequency mixing (in the shift mixer); the incoming signal, converted to intermediate frequency by mixing with the local oscillator (in the signal mixer), is amplified and mixed with the common oscillator frequency (in the output mixer). By choosing the appropriate sidebands of the mixing process, microwave-oscillator frequency drift is cancelled out, and the output and input frequencies differ by exactly the shift frequency. Microwave-oscillator drift modifies the intermediate frequency, which is less serious. Interconnecting the transmitter and receiver in this way introduces several possible sources of feedback, and filters in a number of paths are necessary. Detailed analysis of the problem is beyond the scope of the

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Mr. Craven and Mr. Knight are with Standard Telecommunication Laboratories Ltd.

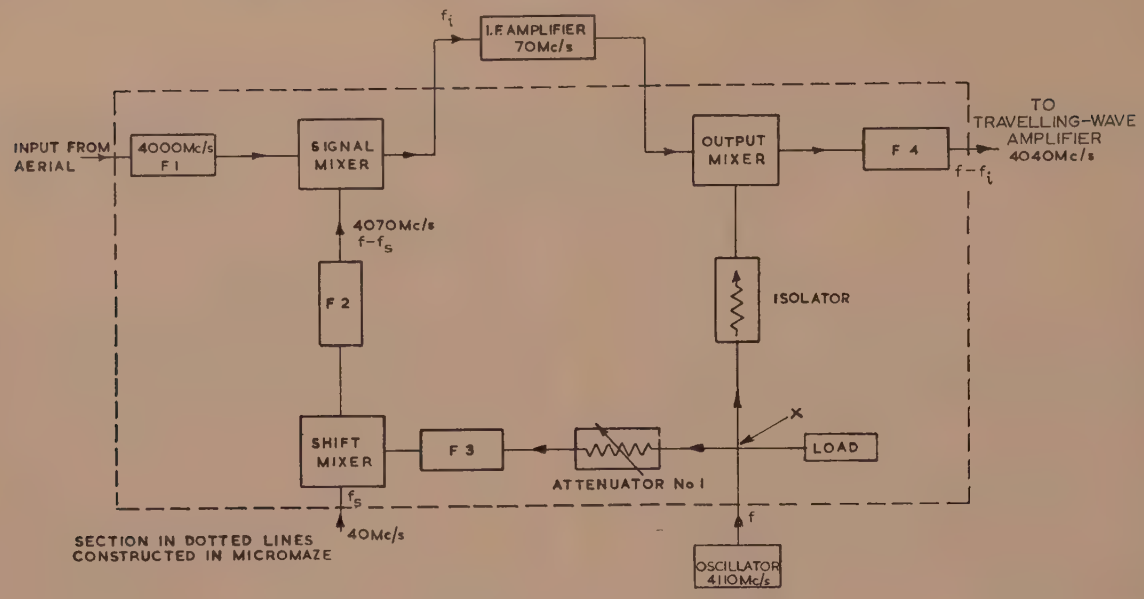


Fig. 1.—Block schematic of microwave repeater.

present paper, but the main purpose of the various filters can be briefly outlined:[5]

(*a*) Denoting the common oscillator frequency by f, the shift frequency by f_s, and the intermediate frequency by f_i, we have, referring to Fig. 1, the possibility of $f - f_i$, generated in the output mixer, travelling to the shift mixer where it is converted to $f - f_i - f_s$, i.e. it is equal to the signal-input frequency. This feedback produces echo distortion and must be highly attenuated. The necessary attenuation is partly supplied by the directive coupler, X, and the two balanced mixers: the remainder, about 90 dB, must be supplied by filters F2 and F3.

(*b*) Another possibility is that the frequency f passing through the shift mixer mixes, at the signal mixer, with the adjacent channel (4 000 Mc/s in Fig. 1) received through the aerial circuit—this produces the intermediate frequency f_i. About 70 dB suppression is required and this is provided by F1 and F2.

(*c*) The shift mixer produces the sidebands $f \pm f_s$ which must be prevented from reaching the output mixer where conversion of $f - f_s$ to $f - f_s - f_i$ occurs. This is equal to the signal-input frequency, and can be transferred by aerial back-coupling to the input circuit. Attenuation through the aerial back-coupling path will not normally be sufficient, and extra suppression must be provided by F3 and F4.

From the foregoing it is evident that a number of unwanted frequencies are generated, and their elimination from critical parts of the system is a relatively complex problem. The net saving when the system is compared with one employing automatic frequency control is nevertheless very great. The stability required by the microwave oscillator in the present system is dictated not by the need to restrict drift in the emitted frequency—since this is effectively independent of the microwave-oscillator frequency—but by the necessity of centring the intermediate frequency on the i.f. amplifier pass band. This is a much less stringent specification and is easily met by mounting the oscillator cavity in a temperature-controlled oven.

(3) BASIC CONSTRUCTION

The most convenient type of construction for any complex assembly is likely to be of planar form in which the waveguide routes preferably run side by side. The final shape taken will depend not only on the waveguide circuitry, but also on the type of components available and the space into which the assembly must be mounted. In any arrangement which involves a number of waveguide routes, the final mechanical design is inevitably influenced to a major degree by the type of multi-port components (such as hybrids) chosen. Where these components are necessary, emphasis on simply constructed designs is essential wherever possible.

The microwave repeater of Fig. 1 requires two balanced mixers, and normally these are realized by mounting crystals in a hybrid-T. However, in many applications a 3 dB directional coupler is a satisfactory substitute for the hybrid-T. It is not only simpler to make, but helps greatly in achieving a simple waveguide-route layout. Directional couplers of this type employ slot coupling through the guide walls. Power transfer between the two guides forming the coupler may be obtained by using either the broad or the narrow dimension as the common wall:[6,7] the former then gives a design in which the waveguide is stacked on the broad face. Thus, if there are more than two guide routes, the inner guides become inaccessible through the broad face. This is inconvenient for crystal mounts or where filters are to be tuned, so that in many applications this style of construction is unattractive.

The general design that results when the hybrid-T's are replaced by 3 dB directional couplers and the directional coupler interconnecting the transmitter and receiver sections (X in Fig. 1), all employing common narrow-wall coupling, is shown in Fig. 2. The somewhat complex route followed by the various frequencies is indicated, and because of its obvious resemblance to a maze the term 'micromaze' has been adopted. The construction can readily be achieved by soldering sections of standard waveguide together to form the necessary routes—an example of this kind of construction will be shown later. However, there are possible advantages in a more radical approach. Since micromaze forms a continuous surface the whole unit can be constructed from brass or copper sheet. Made in this way, channels can be milled in the base plate and the waveguide walls inserted into them. The various elements that go inside the guide, such as coaxial-transformer parts and loads, are then installed. The top plate, with wall channels corresponding to the base plate, is then mounted in position and the construction completed. For certain applications, the walls may be held in by screws and adequate isolation between sections achieved. However, for minimum leakage, as in the present example,

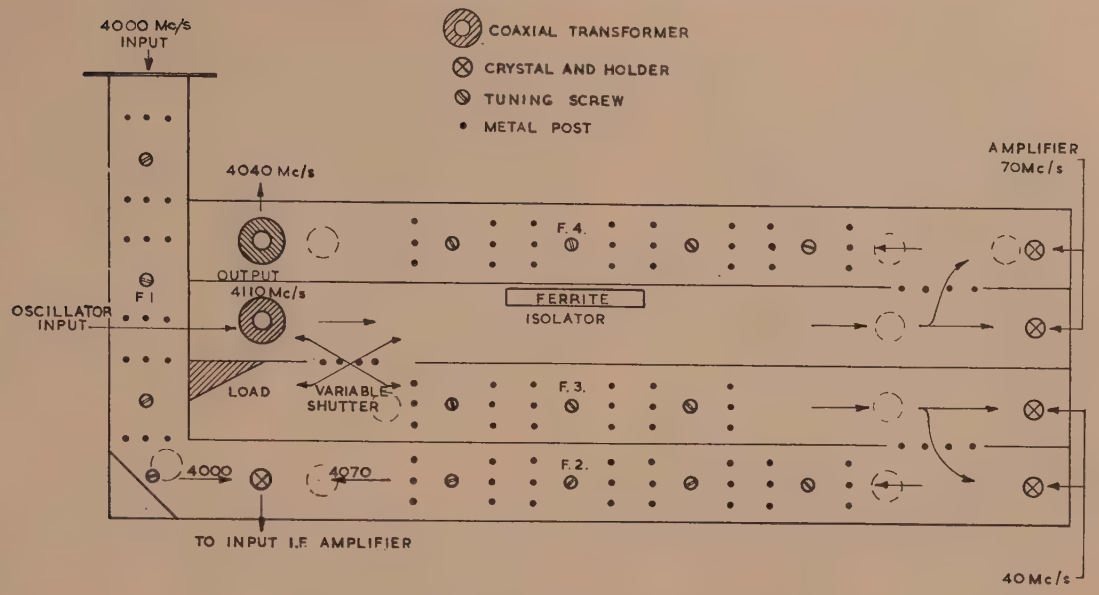


Fig. 2.—Microwave repeater in micromaze construction.

soldering is essential. This can be achieved by lightly tinning the wall contact surfaces and channels before assembly. The complete unit is then placed in an oven and raised to the necessary temperature. This style of construction is well suited to the comparatively large waveguide assemblies needed in the 7·5 cm band. At shorter wavelengths, if manufacture in large quantities is necessary, there is justification for considering die-casting techniques.

(4) DESCRIPTION OF COMPONENTS

The various components required in a repeater are well known but not all can be adapted to integral assemblies. In this Section the most suitable types are briefly reviewed.

(4.1) The 3 dB Coupler

A 3 dB directional coupler is a satisfactory substitute for a hybrid-T in many applications, particularly in balanced mixers. Planar types employing coupling through the common narrow wall have been described.[7,8] The directivity is adequate, and the bandwidth for equal power division is reasonably broad. In general, the bandwidth of Riblet's design[8] is greater than that of the multiple-post-type coupler.[7] However, the latter is easily designed and, where necessary, power transfer can be modified simply by changing post diameter. Equations giving power transfer as a function of the various parameters are derived in the original Reference. Expressed in terms of slot length and the reciprocals of the guide wavelengths of the two propagating modes, the power transfer is given by

$$P_3/P_1 = \sin^2\pi l_c\left\{[1/\lambda^2 - (X/\pi\lambda_c)^2]^{1/2} - [1/\lambda^2 - 1/\lambda_c^2]^{1/2}\right\} \quad . \quad (1)$$

where $$(\tan X)/X = (-l/\pi a)[\log_\varepsilon (l/2\pi r) + \mathrm{F}(\theta, l/\lambda)] \quad . \quad (2)$$

The correction term, $\mathrm{F}(\theta, l/\lambda)$, is normally small and for approximate design can be neglected. Where high accuracy is required, the angle of incidence of the H_{01} wave on the grid of posts is needed:

$$\cos\theta = X\lambda/2\pi a \; . \; . \; . \; . \; . \; . \quad (3)$$

The correction for the angle of incidence is obtained graphically in Reference 9, or in the earlier Reference 7. Fig. 3 illustrates a typical performance and the degree of agreement with theory obtained by the exact method. An approximate expression for bandwidth is derived in Appendix 12.2.

A short slot is usually desirable, and this may be obtained by using large spacings between posts (but not exceeding $\lambda/2$) of small diameter. Typical slot lengths in the 7·5 cm band are between 4 and 6 in.

(4.2) Balanced Modulator

If two crystal diodes are added to a 3 dB directional coupler and the phasing is suitably arranged, a balanced mixer is obtained. A 90° phase lag occurs in the coupling region, and for this reason the crystal mount in the secondary arm must be advanced by $\lambda_g/4$ with respect to the other if an equivalent to the conventional hybrid-T balanced mixer is required (see Fig. 4). An alternative arrangement is shown in Fig. 5. This operates in the following way. As a result of the coupling-region phase lag, the crystals are excited in quadrature. Modulating the crystals in parallel then produces identical sidebands in quadrature, which add in arm 4 because both waves have passed once through the coupling region, but cancel in arm 1 because one wave only has passed through the coupling region and then returned. The second arrangement, which is more convenient for construction and results in a simpler modulation connection, is basically different from the conventional balanced mixer, but both have the same effect of transferring the sidebands to arm 4. The response to second-order effects, such as generated image, is different. This feature of transferring the output components to another arm is very valuable, for it permits the direction of travel to be reversed. Reference to Fig. 2, where the shift and output mixers are of this type, illustrates this. The crystals used in these mixers were experimental germanium diodes matched to the guide through coaxial transformers.

(4.3) Filters

Much literature on microwave filters exists, details of which can be found elsewhere.[10–12] There are two basic types of microwave filter—direct-coupled and quarter-wave-coupled. Either can be used (the equivalence of the two is shown in Reference 11), but present-day opinion is not unanimous on their relative merits, and much probably depends on the application. 'Quarter-wave coupling' is a general term describing filters in which the

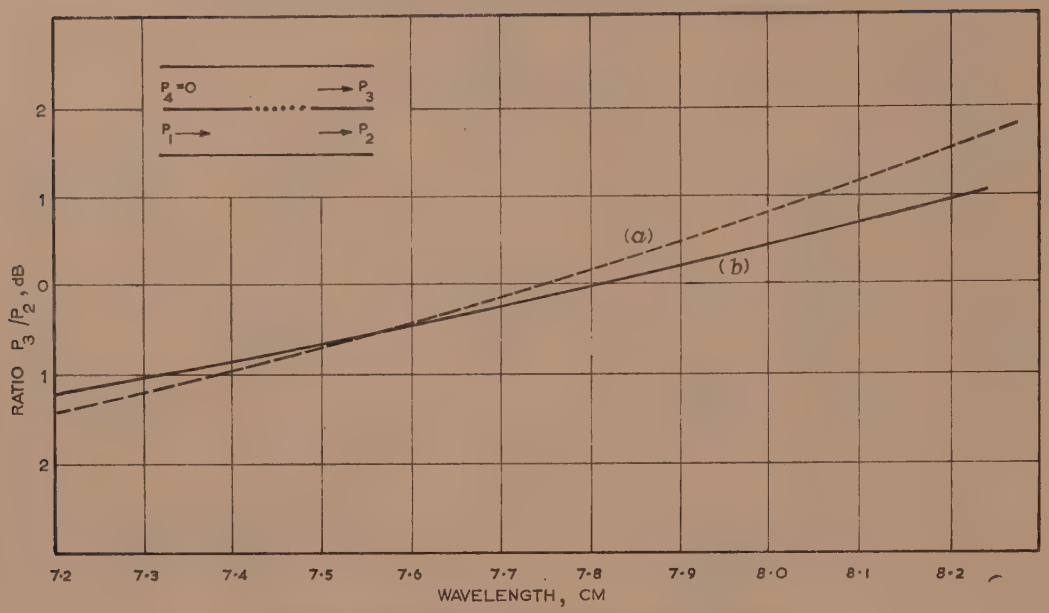
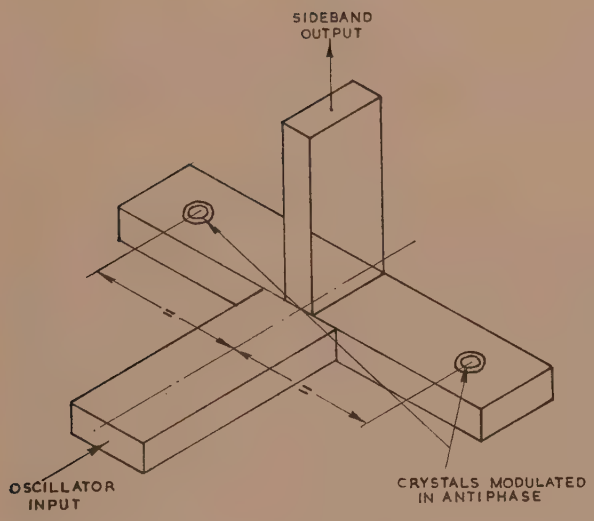


Fig. 3.—Power-division characteristic in 3 dB coupler.

----- Calculated.
—— Measured.



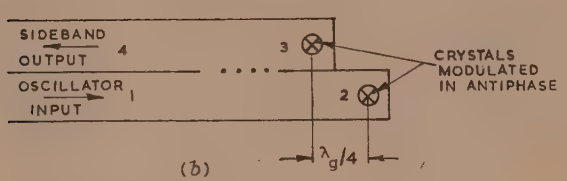


Fig. 4.—Hybrid-T and equivalent 3 dB hybrid.

cavities are connected by guide an odd multiple of a quarter-wavelength long. Filters employing the simple post, or iris, type of susceptances suffer from higher-order mode interference unless three-quarter-wave couplings are employed, and more complex obstacles are needed if the shorter coupling is used.[13,17] The comparison is thus between three types: direct-coupled,

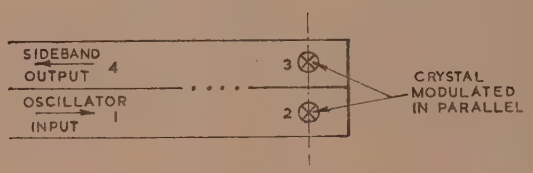


Fig. 5.—Mixer with 3 dB hybrids.

three-quarter-wave-coupled with simple obstacles, and quarter wave coupled with more complex obstacles.

The three-quarter-wave-coupled filter is the easiest to make. Inductive posts are usually preferred, and these become quite thick for large susceptances. They then cease to behave as pure shunt susceptances, and unless this is taken into account the performance in the pass band is somewhat degraded. This is the longest filter of the three.

Quarter-wave-coupled filters usually employ three posts equally spaced across the guide width. The posts are quite thin and behave as substantially pure shunt susceptances, unless very high susceptances are required, in which case five posts can be employed. The use of a number of thin posts makes construction more difficult. The length is approximately 55% of that of the three-quarter-wave-coupled filter.

The direct-coupled filter is the shortest of the three, being about 45% of that of the first type, and also has the advantage of having slightly lower mid-band insertion loss. The principal disadvantage is that even moderate Q-factors call for very high susceptances. Five posts, or more, per susceptance are required and greater care in measurement is needed.

One of the factors influencing the choice of filter in the past has been the ease with which it can be tuned. This has frequently caused a preference for the quarter-wave-coupled type. The non-resonant connecting sections permit the filter to be broken up into its separate resonant cavities which can then be independently tuned before assembling the complete filter. Obviously this does not have any direct application to an integrally constructed assembly, but it is, nevertheless, a great advantage to have non-resonant lengths between the cavities.

Because of its complexity, manufacturing faults must sometimes be expected. If, for instance, a dry joint exists in one of the inductive posts of a multi-section filter it will ordinarily be difficult to find. However, with quarter- or three-quarter-wave couplings it is possible to place directional-coupler holes between cavities. A waveguide arm can then be coupled externally to form a directional coupler. One or two holes suitably located in a filter then permit the source of trouble to be localized more easily. This aspect will be discussed in greater detail in Section 5.

It is evident that in the present application there are advantages in the spaced resonant-cavity type of filter. However, it was felt that the standard type with three-quarter-wavelength spacing was too long, particularly at the lower frequencies, and so the quarter-wave coupling, with the triplet posts referred to earlier, were developed for the purpose. The theory and application has been treated elsewhere[13] and will not be discussed in the present paper.

(4.4) Signal Mixer

The signal mixer is based on a conventional type of crystal mount using a coaxial transformer to match the crystal to the guide. Crystal mounts normally terminate the guide with a short-circuit plate located approximately a quarter guide-wavelength behind the crystal. In micromaze construction a somewhat different arrangement is advantageous. An example is shown, in Fig. 2, in which the local oscillator and signal are incident on the crystal from opposite directions, each circuit containing a filter. F1 is transparent to the signal frequency but rejects that of the local oscillator. Thus, by appropriately locating F1 relative to the crystal, the filter serves as a 'back-plate' at the local-oscillator frequency. In the same way F2 serves this purpose at the signal frequency.

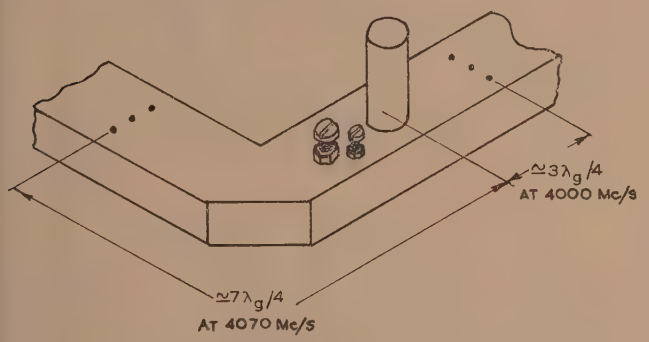


Fig. 6.—Signal mixer.

A diagram of a complete mixer is shown in Fig. 6. It includes a capacitive screw for spot matching the crystal, and an image-rejection screw, tuned to the generated image, which avoids the more frequency-sensitive arrangement that occurs when F1, which is nearly two wavelengths distant, reflects the image back to the crystal.

(4.5) Variable Directional Coupler

The functions of a variable attenuator and directional coupler are combined in a variable directional coupler (X in Fig. 1). The coupler is of the same basic type as the 3 dB couplers, the maximum coupling being approximately 9 dB. Variable coupling is obtained by a sliding shutter, made of beryllium-copper, which varies the effective slot length. Pressure contact on the wall is maintained by contact dimples in the shutter, which slides in channels at the top and bottom of the guide. Other schemes are possible. Variable power transfer can be obtained with non-contact devices. The coupling is sensitive to guide wavelength, and if this is varied by moving a capacitive plate

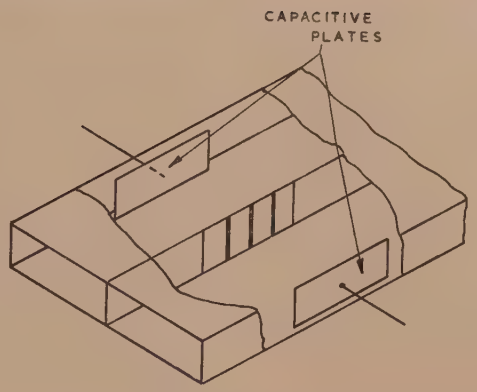


Fig. 7.—Non-contact variable coupler.

in one or both guides, changes up to 10 dB in transfer can be obtained. The arrangement, which is illustrated in Fig. 7, introduces a ridge into the guide, thereby decreasing the guide wavelength and, consequently, the power transfer.

(4.6) The Ferrite Isolator

The ferrite isolator has a dual purpose. It isolates the oscillator from the pulling effect associated with output-mixer mismatch, and prevents sidebands, generated in this mixer, being reflected back into it, after suffering undesirable phase delay, by the oscillator. The isolator is a resonance-absorption type and was an early experimental model.

(4.7) Complete Integral Assembly

The complete integral assembly is shown in Fig. 8. The design realizes the block schematic of Fig. 1 in a practical form of microwave plumbing, and illustrates the way in which the problem of integral design has been solved. Although it is not intended to be exactly representative of current repeater-design practice, a realistic approach has been adopted. For instance, the filters conform, in general, to the specifications outlined in Section 2. The shift frequency employed is not indicative of modern practice, and normally one would expect the output filter to terminate in a travelling-wave amplifier. However, it is felt that the most important consideration is to demonstrate that such an assembly can be adjusted to the high standard of microwave performance—i.e. low reflection and high isolation—required by microwave links. Discussion of the method and apparatus needed to achieve this comprises the remainder of the paper.

(5) ALIGNMENT AND TESTING TECHNIQUE

(5.1) Introduction

The general nature of the problem is evident from Fig. 8. Clearly a tuning and testing technique differing radically from the conventional slotted-line method is necessary. The main points of entry to the assembly are the input and output ports, but additional ports that may be useful are offered by the crystal mounts. Otherwise, the unit must be considered as sealed, so that most of the components are electrically inaccessible. In providing a solution to these difficulties, the present example was regarded as an essay in a principle of wider application, and therefore an equally general alignment technique was sought.

(5.2) Choice of Directional Coupler

It can be seen that the signal input connection is the only port which does not involve entry through a junction which itself may have to be matched. Thus additional entry ports must

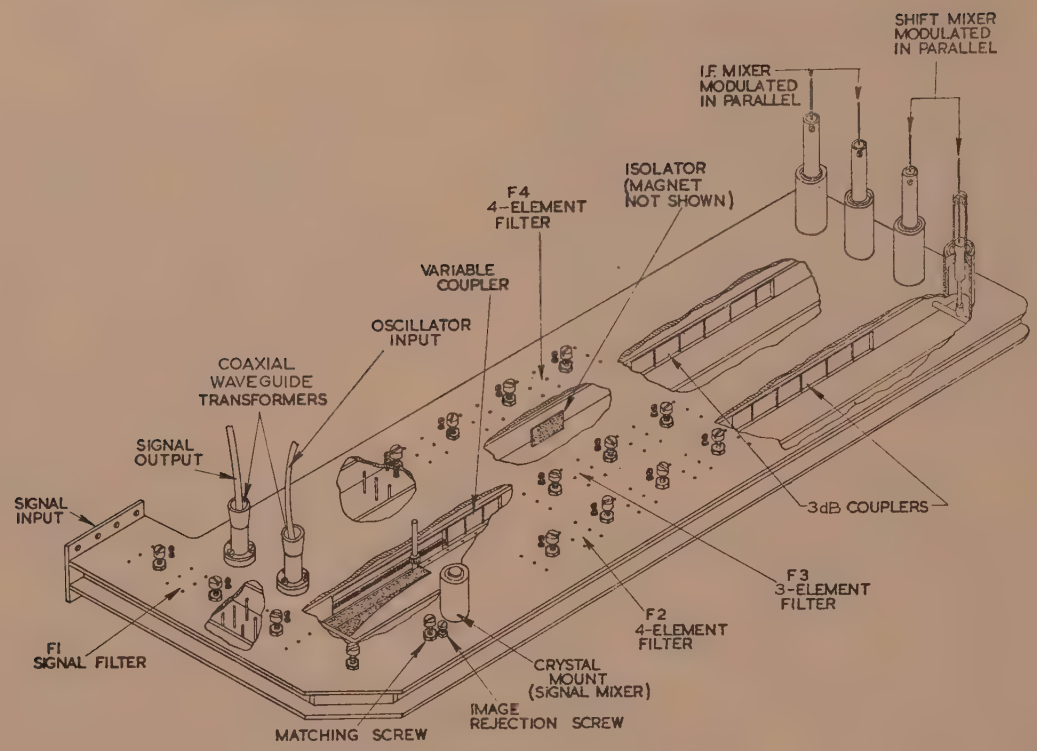


Fig. 8.—General arrangement of micromaze repeater.

be provided. General requirements on these ports are that they must have no effect on the electrical performance of the apparatus, be simple to make, and permit measurement of reflection. These specifications can be almost ideally met by a Bethe hole coupler—a directional coupler having a hole in the centre of the guide.[14,15] The additional waveguide arms can be formed by coupling a suitably terminated waveguide section to the hole as shown in Fig. 9. The rectified output of this arm, when amplified, may be used to indicate reflection, and the waveguide arm may then be treated as a test piece to be connected to any hole from which information is desired. The virtue of the Bethe coupler in this application lies in its simplicity: a circular hole in the centre of the guide is easily made and can be plugged up when not in use, and directivity is adjustable by the angle at which the guides intersect. Thus a common size of hole can be used at all frequencies, the adjustment for maximum directivity being obtained, mainly, by modifying the angle between the guides.

The general principle of alignment that applies in a complex example where a number of networks are in cascade is as follows. Non-dissipative networks, such as filters or transducers, normally terminate in some form of matched load. In a standard test-bench arrangement the termination is a dummy load designed for the purpose, but in a practical system the load is, typically, a crystal diode or the input of a travelling-wave amplifier. This load must be matched for zero reflection, and a directional-coupler hole preceding it, as in Fig. 9, becomes necessary. In this arrangement a matching transformer giving continuous variation of both magnitude and phase is most convenient. With the matched load provided, the waveguide test piece is moved to the next coupling hole and the directional coupler so formed used to indicate the reflection of the lossless network immediately preceding the load—a filter, in the example. This is then adjusted for zero reflection, and the process repeated at the input to each successive network until the source is reached. This example assumes that the networks are at least approxi-

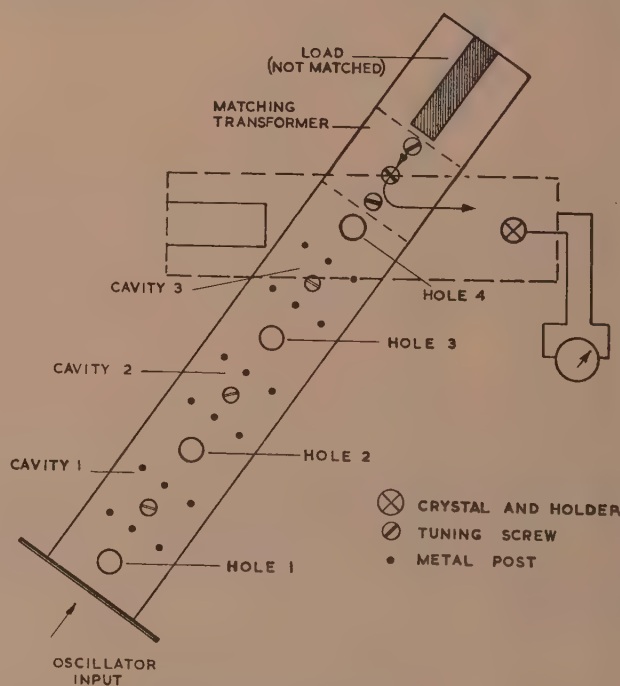


Fig. 9.—Basic principle of alignment.

mately transparent initially, a condition that could easily be obtained.

(5.3) Swept-Frequency Methods

Quarter-wave-coupled filters have been satisfactorily tuned by the method outlined in the previous Section, but, nevertheless, this method has serious practical disadvantages when applied to sharply resonant networks. Filter characteristics over a band

of frequencies are normally of interest and measurement is tedious. Furthermore, with the foregoing technique applied to a multi-section filter, a progressive deterioration in the match, or, alternatively, correction at a later stage in a highly frequency-sensitive manner, is possible. If the latter occurs, tuning must be repeated and the response characteristic measured again. This can all be avoided if a swept-frequency display giving a continuous presentation of results is used.

In recent years, swept-frequency methods have gained considerable favour in microwave applications. However, their application in the present instance is not so simple because of the power levels involved. This is a result of the small reflections, low signal powers and attenuation of coupling holes, which, in the Bethe coupler, is very high if the hole is small. A number of large holes in an assembly solely for test purposes is obviously objectionable, and, in any case, space may not always be available. A hole of $\frac{5}{8}$ in diameter in a 2 in × $\frac{3}{4}$ in guide has an attenuation of about 40 dB when wall thickness is taken into consideration. For accurate matching a 1% voltage reflection should be measurable, i.e. a signal 40 dB down on the incident power. A typical value for the latter is 100 mW, so that the power available for the display equipment is 10^{-9} watt. This is favourable because the power in the guide must frequently be restricted to the maximum permitted by some device such as a mixer crystal. An example of this occurs in the signal mixer. Local-oscillator power at this crystal will be about 10^{-3} watt, so that it is reasonable to restrict signal power to about 10^{-5} watt. Under these circumstances, with a 40 dB coupling-hole loss, the power available from a 1% reflection for display purposes will be 10^{-13} watt. A more favourable arrangement, which is discussed in the next Section, improves this to 10^{-9} watt. Thus, allowing a 10 dB safety factor, the sensitivity required by the display receiver is −70 dBm. For this reason it was decided to develop a high-sensitivity display receiver with a minimum sensitivity of the above specification. Preferably, higher sensitivities should be available in order to cope with the attenuation of smaller holes, or generally more unfavourable conditions. The design of this specialized test equipment is discussed in a later Section.

(5.4) Alignment of Signal Input Channel

The signal input channel includes the input path proper (consisting of F1, the right-angle bend, crystal-matching screw, image-rejection screw and crystal) and the local oscillator path (consisting of F2, the shift mixer and F3). In describing the alignment procedure it is assumed that a swept-frequency display of adequate sensitivity, giving either reflection-magnitude or transmission-amplitude response (in Cartesian co-ordinates), is employed. A directivity of better than 40 dB is assumed in the directional coupler.

(5.4.1) Tuning F3.

The filter F3 of Fig. 10(*a*) is tuned with its mid-frequency (4 110 Mc/s in the present example) located at the common oscillator frequency. With the sweep oscillator centred at 4 110 Mc/s, a sweep width of about 60 Mc/s and the output connected to the oscillator input transducer, the coupling arm is mounted at hole No. 5 and set so that it receives the incident wave (adjusted to 180° different from the reflection setting). A moderate load (v.s.w.r., 0·7), which can be obtained by a matched load plugged into the crystal mount or by suitably biasing the crystals, terminates F3. Initially, all the main-tuning screws are withdrawn from the guide, with a few threads of the fine-tuning screws inserted. With the display receiver set for maximum sensitivity (the attenuation of all the cavities that are off tune will be high), a screw is introduced into any of the cavities until a ripple appears on the oscilloscope screen. The procedure is repeated until a rough response curve is obtained. More careful adjustment, first of the coarse screws and finally of the fine-tuning screws on each cavity, is then made until a satisfactory flat pass band is obtained (the filters are designed for a maximally flat response). During the alignment, the frequency should be checked by a wavemeter (showing a resonance dip in the display) in order to ensure that the response is correctly centred. If desired, final adjustment can be obtained by mounting the test guide at hole No. 6 (set for reflection) and tuning for a smooth reflection curve of a magnitude consistent with the filter termination employed. However, in general, a flat amplitude response is sufficient for filters in the local-oscillator path.

(5.4.2) Tuning F2.

F2 may be tuned in a variety of ways: one method is that illustrated in Fig. 10(*b*). The output of the swept-frequency oscillator is connected to the test guide arm which is mounted at hole No. 3 and set to couple power directively into F2. The oscillator sweep is centred at 4 070 Mc/s. The display receiver is connected to the shift-mixer crystal mount, one of the loads from the previous tests being retained. The screws on F2 are then adjusted in the manner described in the previous Section. It is sometimes possible for F1, which is not yet tuned, to resonate near the frequency of F2, causing ripples at the band edges of F2. This can be eliminated by introducing the tuning screws a few turns into the cavities of F1. Reflective interaction at the preliminary tuning stage is always possible with adjacent filters, but disappears when the filters are even approximately tuned. The match of the shift-mixer crystals is not particularly critical and as the mount is broad-band no adjustment is required. Coupling holes are provided for checking the match.

(5.4.3) Tuning F1 and Adjusting Crystal Match.

The crystal and filter F1, shown in Figs. 10(*c*) and (*d*), are in the signal path and must be adjusted for a very low reflection. Initially the power is directionally coupled into F1, connection to the swept-frequency display apparatus being made at the signal input port. With an appropriate sweep width, the input filter F1 is tuned for a good transmission response. With this procedure completed, the direction of coupling is reversed and power directively coupled towards the crystal; the display apparatus then receives reflection. This arrangement is preferable to the reciprocal one discussed in Section 5.3 because the sensitivity needed in the display apparatus is reduced by about 40 dB. The attenuation of the coupling hole reduces the power at the crystal to the desired power of about 10^{-5} watt, and no further attenuation is necessary.

The mixer crystal is terminated in its normal circuit, except that the crystal current is monitored by a suitable milliammeter. Local-oscillator power is obtained from the microwave oscillator by shift-mixer generation, the level being set at the desired value by the variable directional coupler (X of Fig. 1). The crystal match is more conveniently adjusted with the sweep motor stationary, and therefore the oscillator is merely set at the signal frequency (4 000 Mc/s here). The crystal-matching transformer is then adjusted for the best match obtainable, which should be about 1% voltage reflection.

The final step is to couple power directively into F1 through hole No. 1. With the sweep motor rotating, the reflection curve of the complete input section at signal frequency is displayed. This is principally the curve of F1 but it also includes additional reflections, such as those occurring at the corner, and any frequency sensitivity in the crystal match. The fine-tuning screws on F1 are then adjusted for a smooth curve of low reflection at the centre. Ripples in the curve are always possible because of

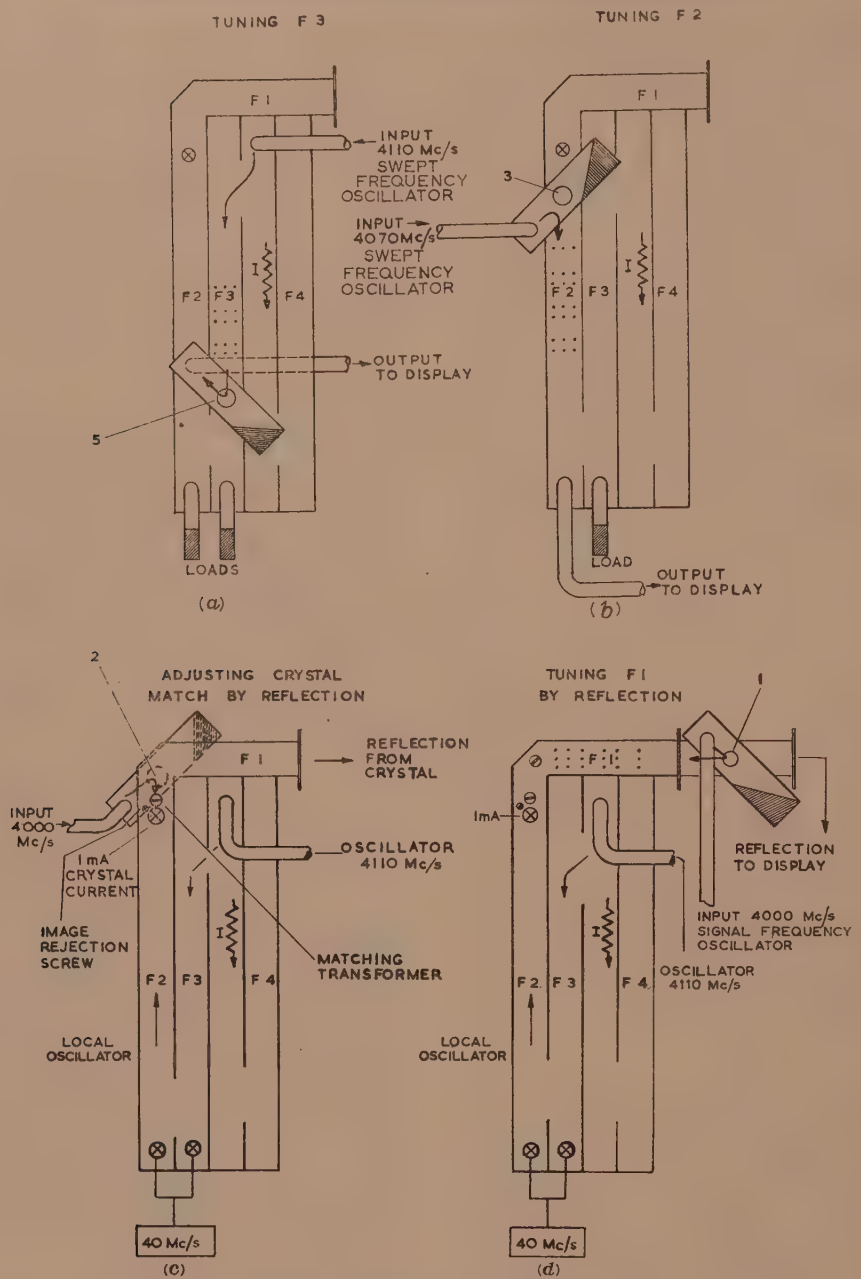


Fig. 10.—Alignment of sections.

(*a*) and (*b*) Local oscillator. Directional coupler set for transmission response.
(*c*) and (*d*) Signal input.

the generated image, reflected back by F1, upsetting the match in the pass band. There was evidence of this in the present design; an image rejection screw, located adjacent to the crystal, was therefore introduced and adjusted until the ripple disappeared. This affected the crystal match, which was then readjusted to the previous value. At this stage, small adjustments to the corner screw can be made, together with minor readjustments to the crystal match and fine-tuning screws.

(5.5) Signal Output Channel

A similar technique is employed in the alignment of the output channel, and so it will not be described in detail. The tuning adjustments are materially less complicated, since only one filter, F4, is involved.

(6) SETTING UP THE DIRECTIONAL COUPLER

The directivity of the directional coupler used for testing limits the ultimate match obtainable; for a 1% reflection to be detectable, a directivity of 40 dB or better is necessary. The directivity is controlled mainly by the angle between the two guides, and is also a function of frequency. However, for maximum directivity, a small additional reflection must be introduced and adjusted to cancel the residual coupling remaining at the optimum angular setting. A simple capacitive screw serves the purpose, although a calibrated probe adjustment would be preferable for regular use. The angle between the coupled and main guides must be accurately maintained, and a jig is necessary to hold the coupled guide in position on the main assembly (shown in Fig. 11). When mounted in position the jig

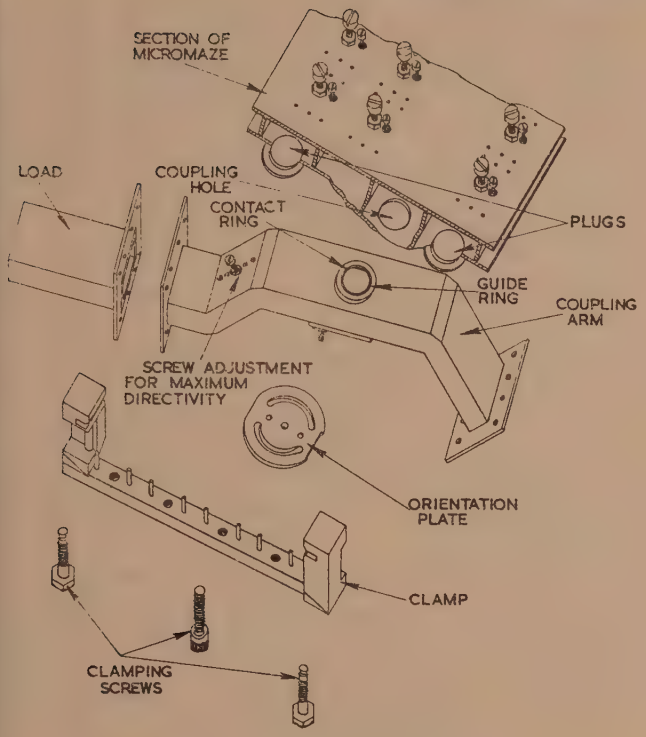


Fig. 11.—Coupler mounting jig.

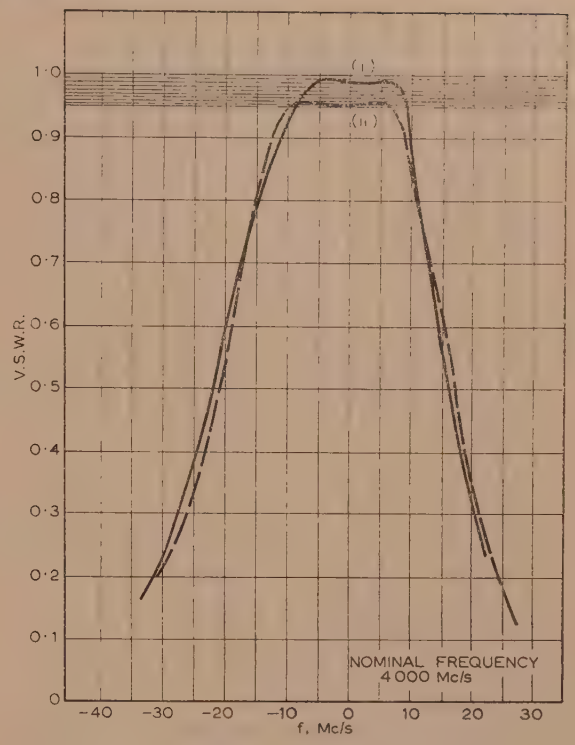


Fig. 12A.—Performance of signal-input-channel v.s.w.r.

holds the contact ring of the coupled arm firmly in contact with the main guide; location with respect to the hole is obtained by an annular recess cut in the guide wall. The angular setting for any given frequency is fixed by locking the two nuts holding the rotatable plate. For frequent use a calibrated dial would preferably be included on the plate. Although not particularly elegant, the arrangement was quite effective for laboratory use: for rapid use in production a more easily mounted device would be desirable.

The angular setting and capacitive screw adjustment for a given frequency is determined by adjusting the directivity on a test bench. A section of guide, which duplicates the hole and guide dimensions of the main assembly, is employed as the main arm of the directional coupler. The main arm is then terminated in an accurately matched load; the coupled arm is swung for maximum directivity, the final adjustment being made with a capacitive screw. With suitable arrangements, calibration of settings for a range of frequencies can be obtained, and the device set up quickly at any frequency. Mechanical tolerances are obviously important in such a method, and their effect on directivity is analysed in Appendix 12.1.

(7) STANDARD OF PERFORMANCE

The method of alignment outlined in preceding Sections is regarded as rather novel, at least as applied to an integral assembly. It is therefore reasonable, for comparative purposes, to relate the performance to that obtainable by conventional techniques. For this reason, after the tuning procedure described had been carried out, the results were checked using standard methods: the v.s.w.r. was measured with a slotted line, and calibrated attenuators were employed to determine the insertion-loss curves. It is the latter measurements, in preference to swept-frequency oscillograms, that are presented to show the performances obtainable.

The performance of the signal input channel is shown in Figs. 12A, 12B, 13A and 13B. The v.s.w.r. was measured with

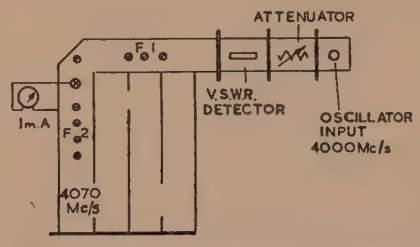


Fig. 12B.—Measurement arrangement for the input v.s.w.r.

the slotted line mounted at the input flange (as shown in Fig. 12B), the signal level being maintained at a suitably low level so as not to affect the match. Two curves are shown: curve (i) probably represents the very best obtainable and was achieved after the very careful adjustment described in Section 5.4; curve (ii) shows what can be obtained with slightly less care and should be achieved in practice with only moderate experience in the method. The agreement between these curves and the swept-frequency display (bearing in mind that one is a plot of reflection characteristic, and the other represents v.s.w.r.) was good. The difference between curves (i) and (ii) was readily detectable, provided that the low-reflection part of the curve was displayed on an expanded scale.

The corresponding curves can be obtained for the signal output channel. The results, which are not materially different from the input channel, are summarized in Table 1.

Table 1

SUMMARY OF RESULTS

Input channel.—Reflection, 2% over 21 Mc/s band; bandwidth between 3 dB points, 56 Mc/s.

Output channel.—Reflection, 2% over 22·5 Mc/s band; bandwidth between 3 dB points, 56 Mc/s.

Local-oscillator sections.—

Bandwidth of F_3 and associated components, 42 Mc/s.

Bandwidth of F_2 and associated components, 32 Mc/s.

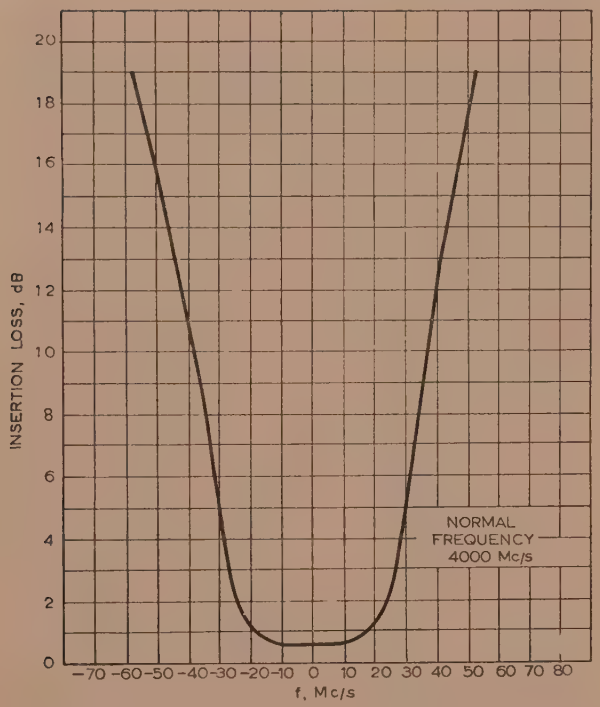


Fig. 13A.—Performance of signal-input-channel insertion loss.

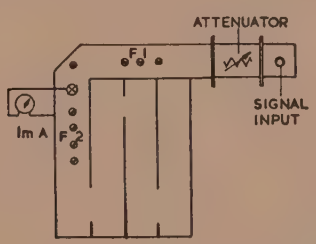


Fig. 13B.—Measurement arrangement for the input insertion loss.
Curve measured for constant signal at 1 mA crystal current. No local oscillator.

The transmission characteristic of F2 is shown in Figs. 14A and 14B. The performance of F3 is similar.

(8) SWEPT-FREQUENCY RECEIVER DESIGN

It is evident that specialized test equipment is essential for the *in situ* alignment technique. General specifications that can be defined are a useful display output for an input of −70 dBm and a frequency sweep of ±50 Mc/s about the centre frequency with a substantially constant output. The required performance should be available throughout the useful band (3 600–4 200 Mc/s in the example). In addition, a linear law of presentation is preferred to a square law.

The desired amplification can be obtained in a number of ways, but if high sensitivity is essential radio-frequency amplification is needed. The system chosen is basically a superheterodyne receiver in which the local oscillator is derived from the sweep oscillator by frequency mixing (Fig. 15). This scheme, which has much in common with the shift-frequency system in the repeater, has the principal advantage that the local oscillator keeps in step with the sweep oscillator. Thus the i.f. bandwidth does not have to accommodate the large bandwidth swept by the oscillator (100 Mc/s) but the much smaller bandwidth required by the information contained in the sweep. Therefore narrow i.f. bandwidths can be employed when a high sensitivity is required.

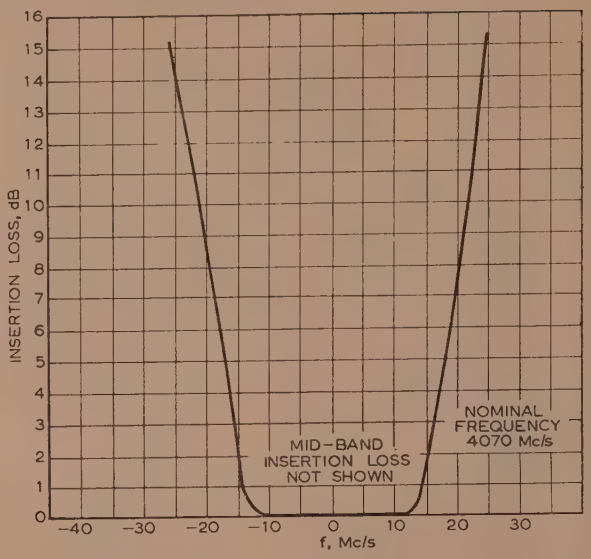


Fig. 14A.—Performance of local-oscillator-path insertion loss.

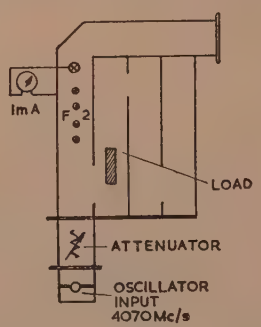


Fig. 14B.—Measurement arrangement for local-oscillator-path insertion loss.
Curve measured for constant signal at 1 mA crystal current.

(8.1) Intermediate Frequency

Although not novel this arrangement presents complications in a high-sensitivity microwave receiver. These problems are inherent in a system in which the local-oscillator signal is generated by a shift-frequency process, and occur if the suppression of the original frequency is incomplete. In the present system we have the swept-frequency oscillator of centre frequency f_0 (Fig. 15) with a frequency sweep of δf_0. When modulated in a non-linear device the customary sidebands appear, each sweeping the band with a frequency $\pm\delta f_0$. If the modulation frequency is f_s the frequency bands covered are: input frequency, $f_0 \pm \delta f_0$; upper sideband, $f_0 \pm \delta f_0 + f_s$; lower sideband, $f_0 \pm \delta f_0 - f_s$.

If the generated local-oscillator frequency (either the lower or upper sideband) is to be correct, f_s must equal the intermediate frequency. Then, unless $f_s > 2\delta f_0$, the upper and lower sidebands overlap the bands swept by the original frequency, $f_0 \pm \delta f_0$, and the undesired frequencies cannot be suppressed by filtering. If $\delta f_0 = 50$ Mc/s, an intermediate frequency exceeding 100 Mc/s is essential. The actual value chosen depends on the practical problems of microwave-filter design. This filter must provide a high attenuation, and since the sweep centre frequency, f_0, may be adjusted for operation within a wide band (3 600–4 200 Mc/s), it must ordinarily be variable. The problems of designing such a filter, together with the disadvantage of an additional tuning adjustment, make a broad-band filter, with

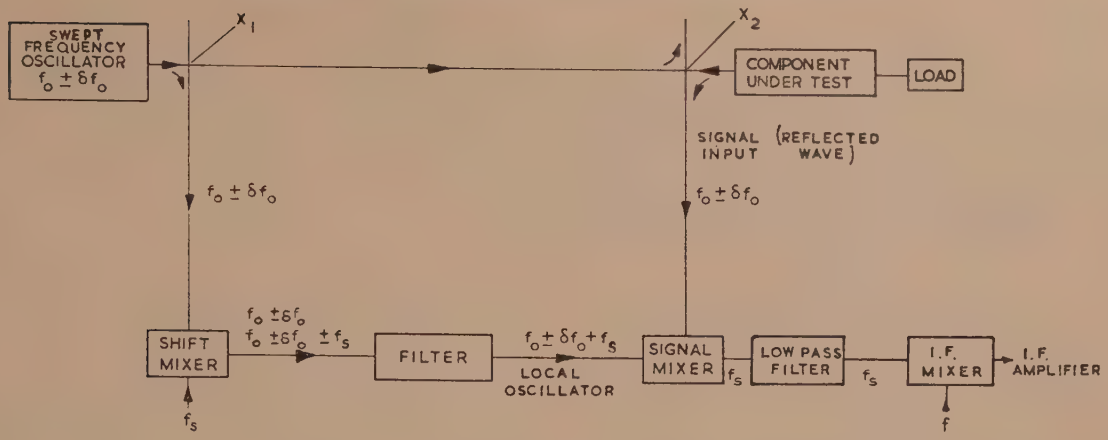


Fig. 15.—Block schematic of swept-frequency receiver.

f_0 = Oscillator centre frequency.
δf_0 = Frequency sweep.
f_s = Shift frequency (1st intermediate frequency).
X = Directive coupler.
$f = f_s - f_{12}$(f_{12} is the 2nd intermediate frequency).

fixed tuning, greatly to be preferred. This is possible if a much higher value than the minimum f_s, just considered, is employed. By similar reasoning it is clear that if $f_s > 600$ Mc/s (operation in the 3 600–4 200 Mc/s band), the original input frequency and two sidebands are all separable by filtering, and the preferred type of filter may be used. For example, if $f_s = 1\,000$ Mc/s, operation anywhere within the desired band gives frequencies in the following ranges: signal input, 3 600–4 200 Mc/s; upper sideband, 4 600–5 200 Mc/s; lower sideband, 2 600–3 200 Mc/s.

Choosing the upper sideband as the local-oscillator frequency, the unwanted frequencies can be rejected by a high-pass filter with a cut-off of about 4 400 Mc/s. This is particularly convenient because the high stop-band attenuation needed can be provided by a simple filter consisting of a section of guide operating beyond the cut-off frequency. The latter requirement is a direct consequence of high sensitivity, and will now be considered.

(8.2) Filter Design

The most important design specification that the filter in the local-oscillator path must satisfy is the attenuation in the signal-frequency band. The signal frequency can reach the signal mixer by one of two possible paths: by reflection from the component under test (i.e. the desired path) or by the local-oscillator path. From Section 5.3, the attenuation in the desired path may be as much as 80–90 dB. The filter attenuation in the signal band must then be equal to this value together with a safety factor of, say, 20 dB. Failure to satisfy the latter requirement makes the receiver unusable: the desired signal is completely swamped if the filter attenuation is much too low, and if both path losses are of the same order ripples appear on the display as a result of mutual cancellation at certain frequencies.

The basic frequencies of the mixing process are the principal sources of spurious responses. However, several harmonics of the 1 000 Mc/s oscillator can be troublesome, and it is necessary to provide a low-pass filter to suppress them. (The most desirable value for f_s is, in fact, about 1 150 Mc/s—a value which prevents any harmonics of this oscillator falling in the pass band of the filter.)

Hence the filter needed for this application is a wide-band-pass filter, asymmetrical in its attenuation characteristics because of the high attenuation needed below the pass band. The design of filters of this type has been considered previously[16] and will not be discussed.

(8.3) The Second Intermediate Frequency

An intermediate frequency in the region of 1 000 Mc/s is inconvenient for amplification and a second frequency change to a more conventional value is employed. A frequency of 940 Mc/s beats with the first intermediate frequency to produce a second intermediate frequency of 60 Mc/s. A standard i.f. amplifier is then used to raise the signal level before detection. A simple coaxial low-pass filter is included in the second i.f. mixer circuit in order to reject harmonics of 940 Mc/s.

(8.4) Components

The two balanced mixers are of the type employing 3 dB hybrids. The third mixer uses a single crystal and is located at the end of the coaxial filter assembly. A monitor circuit directively coupled to the main oscillator input arm permits the oscillator output power to be monitored. An absorption wavemeter in the main arm precedes this coupler and thus provides wavelength indication on both monitor and display apparatus.

(8.5) Construction

Since the resultant microwave assembly is a complex one it seemed worth while to construct it integrally. The result is shown in Fig. 16. In this example conventional manufactured guides were employed, soldered together and bolted at certain points in order to provide adequate strength. The resultant elimination of flange couplings is also an aid in obtaining the high isolation needed.

(8.6) Additional Apparatus

A certain amount of additional apparatus is required. The 940 Mc/s and 1 000 Mc/s oscillators employ disc-seal triodes in coaxial resonant cavities. The swept-frequency oscillator employs a Heil tube in a waveguide cavity, and the centre-frequency is set by the position of the short-circuit plunger, the sweep width being controlled by the penetration of the rotatable blade. A synchronous motor rotates the blade at 1 500 r.p.m. One motor revolution results in four sweep cycles, three of which are suppressed in the display apparatus. An oscilloscope with high gain d.c. amplifiers and facilities for blanking is necessary.

(8.7) Performance

The main object in developing the swept-frequency receiver was to achieve the high sensitivity needed for alignment purposes.

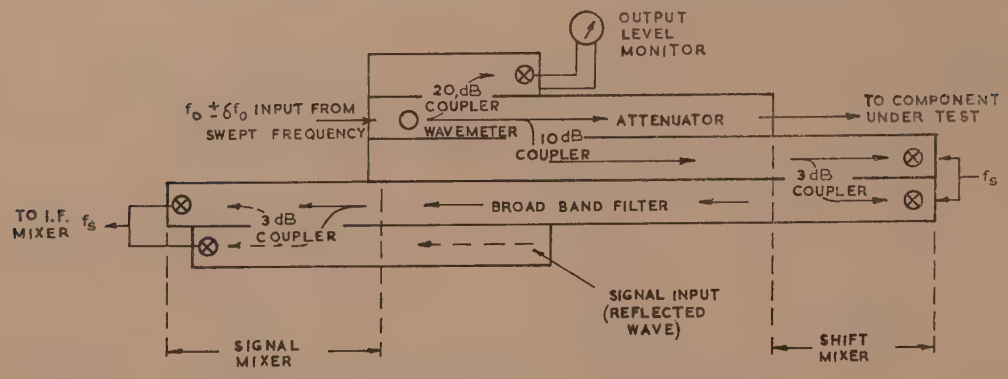


Fig. 16.—Microwave assembly of the swept-frequency receiver.

For this reason only the bare essentials were provided in the display equipment. The display was not directly calibrated, substitution of a calibrated reflection probe in the test arm of the Bethe coupler being employed to determine reflection magnitude. The sensitivity was completely adequate for displaying signals of −70 dBm, and much higher sensitivities could have been obtained, especially if a narrower i.f. bandwidth had been employed. The latter was about 1 Mc/s, the principal factor limiting narrower bandwidths being the stability of the 940 Mc/s and 1 000 Mc/s oscillators. With greater stability, for instance, by employing crystal control, bandwidth could be decreased by a large factor. Some care was necessary in matching the various components in order to prevent amplitude ripples on the display. Variations from this and other causes were approximately 0·25 dB.

(9) CONCLUSIONS

Two examples of the application of integral-construction technique have been given, and it has been shown that this presents great opportunities for space saving. Quantity production of a unit would justify special construction methods, but where only modest quantities are required, fabrication from standard guide is likely to be preferable. The saving in space that is possible can be judged by comparing the microwave repeater in micromaze form with the same design constructed in the conventional manner. The dimensions of the integral design were approximately 28 in × 12 in × 4 in, and the assembly weighed 30 lb. This would fit conveniently in a part of a repeater bay. In conventional construction the same design occupied two repeater bays and weighed about 100 lb.

No claims can be directly made for overall repeater performance, since measurements have been wholly of a microwave nature. From this aspect the performance is considered very good, conventional slotted-line measurements having confirmed swept-frequency-display measurements. The technique of tuning a complete signal circuit as a unit, rather than as separate elements, should eliminate additional sources of reflection and give improved performance. As in all swept-frequency methods of alignment, the technique is fundamentally a rapid one.

The most important disadvantage of the method is that once the waveguide is designed and built, modification is never easy and may be exceedingly difficult. In this respect it represents a challenge to the engineer to investigate all problems completely in the experimental stage and then finalize the design. The microwave repeater shown was largely designed on paper and required few modifications.

(10) ACKNOWLEDGMENTS

The authors wish to thank Mr. H. Grayson, who suggested the scheme, for his encouragement, Mr. L. Lewin for his suggestions, and Mr. G. H. B. Thompson, who developed the ferrite isolator. Integral assemblies posed a number of constructional problems, and these were ably solved by Mr. L. V. Dabbs and Mr. H. J. Casford. Thanks are also due to Standard Telecommunication Laboratories Ltd., for permission and facilities to publish the paper.

(11) REFERENCES

(1) DUKES, J. M. C.: 'The Application of Printed-Circuit Techniques to the Design of Microwave Components', *Proceedings I.E.E.*, Paper No. 2401 R, March, 1958 (**105 B**, p. 155).

(2) GRIEG, D. D., and ENGLEMANN, H. F.: 'Microstrip—A New Transmission Medium for the Kilomegacycle Range', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 1644.

(3) LEWIN, L.: 'Miniaturization of Microwave Assemblies', *Transactions of the Institute of Radio Engineers*, October, 1956, **M.TT-4**, No. 4, p. 261.

(4) JAMIESON, E.: 'Integrated Microwave Circuits', *Electronic Engineering*, 1955, **27**, p. 60.

(5) DAWSON, G., and KORYTKO, T. K. M.: 'S.H.F. Radio Links Using Travelling-Wave Output Amplifiers', *ibid.*, 1958, **30**, p. 276.

(6) MONTGOMERY, C. G.: 'Technique of Microwave Measurements', M.I.T. Radiation Laboratory Series (McGraw-Hill, 1947), Vol. 11, p. 887.

(7) TOMIYASU, K., and COHN, S. B.: 'The Transvar Directional Coupler', *Proceedings of the Institute of Radio Engineers*, 1953, **41**, p. 922.

(8) RIBLET, H. J.: 'The Short-Slot Hybrid Junction', *ibid.*, 1952, **40**, p. 180.

(9) MACFARLANE, G. G.: 'Surface Impedance of an Infinite Parallel-Wire Grid at Oblique Angles of Incidence', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 1523.

(10) MUMFORD, W. W.: 'Maximally Flat Filters in Waveguide', *Bell System Technical Journal*, 1948, **27**, p. 684.

(11) LEVY, R.: 'An Improved Design Procedure for the Multi-Section Generalized Microwave Filter', *Proceedings I.E.E.*, Monograph No. 233 R, April, 1957 (**104 C**, p. 423).

(12) POTOK, M. H. N.: 'The Design of Inductive Post-Type Microwave Filters', *Journal of the British Institution of Radio Engineers*, 1958, **18**, p. 263.

(13) CRAVEN, G., and LEWIN, L.: 'Design of Microwave Filters with Quarter-Wave Couplings', *Proceedings I.E.E.*, Paper No. 2001 R, March, 1956 (**103 B**, p. 173).

(14) SURDIN, M.: 'Directive Couplers in Wave Guides', *Journal I.E.E.*, 1946, **93**, Part IIIA, p. 725.

(15) See Reference 6, p. 858.

(16) CRAVEN, G.: 'Wide-Band Waveguide Filters with Short Linear Tapers', *Proceedings I.E.E.*, Paper No. 2547 R, March, 1958 (**105** B, p. 210).
(17) HORNA, J.: 'Irises for Bandpass Waveguide Filters', *Slaboproudy Obzor*, 1958, **19**, p. 221.

(12) APPENDICES

(12.1) Mechanical Tolerances in the Bethe Coupler

A high directivity in the Bethe coupler is essential if the desired standard of performance is to be obtained. The method of setting up the directional coupler is described in Section 6, and from this it is evident that mechanical tolerances are an important consideration. The effect of the various parameters can be investigated experimentally and checked with the theoretical estimates. The expression for directivity is[15]

$$D = 20 \log_{10} \left[\frac{\cos\theta + \frac{1}{2}(\lambda_g/\lambda)^2 F_E/F_H}{\cos\theta - \frac{1}{2}(\lambda_g/\lambda)^2 F_E/F_H} \right] \quad . \quad . \quad (4)$$

where

$$F_E = \exp\left\{-2\pi t[1/1{\cdot}31d)^2 - (1/\lambda)^2]^{1/2}\right\} \quad . \quad . \quad (5)$$

$$F_H = \exp\left\{-2\pi t[1/1{\cdot}71d)^2 - (1/\lambda)^2]^{1/2}\right\} \quad . \quad . \quad (6)$$

Variations in wall thickness affect the electric and magnetic couplings unequally and thus modify the directivity. The magnitude of the effect can be determined from eqns. (4)–(6). For infinite directivity the denominator of eqn. (4) is zero; a change in wall thickness from the design value then alters F_E/F_H. The wall thickness is not very critical. For instance, if the design thickness of $\frac{1}{8}$in ($\frac{5}{8}$in diameter hole) is increased by 0·030in the directivity decreases to 30dB, whereas experimentally a value of 29dB was obtained. A large error was deliberately chosen in order to make experimental confirmation easier, but practical errors only a fraction of this should be easily obtained. Tolerances on this dimension become more strict for small-diameter holes with thick walls.

The accuracy within which the angular setting must be held, if a given directivity is to be maintained, can be found by substituting values of θ in eqn. (4). In a typical example (with design values: $\theta = 54°$; $\lambda_g = 11{\cdot}1$cm; $\lambda = 7{\cdot}5$cm; $F_E/F_H = 0{\cdot}537$) the directivity drops to 34dB for $\pm 1{\cdot}6°$ from the optimum. Experimentally a tolerance of approximately $\pm 2°$ was obtained. This is considered a relatively large angular error and practical tolerances much less than this figure should be achievable.

The frequency sensitivity can be investigated in much the same way except that the variable in eqn. (4) is now λ_g/λ. For the above example, the bandwidth at the 34dB points is 135Mc/s. This neglects the additional frequency sensitivity caused by the capacitive screw which is introduced into the auxiliary guide in order to cancel minor reflections resulting from imperfections in the load, etc. As a result, the bandwidth was decreased to 33Mc/s, which, however, was adequate for the present application. It is believed that this figure could be improved considerably by careful attention to the match of the various bends and the terminating load employed.

The final parameter affecting directivity is the position of the coupling hole. Nominally in the centre, movement to either side must influence the balance between electric and magnetic couplings. The effect of this can be determined from the equations employed by Surdin.[14] The amplitude coefficients of the H_{01}-mode, excited in the secondary guide, for a hole located anywhere in the common broad face are given by

$$\left.\begin{aligned} A_1 &= \frac{j\gamma_1}{ab}[m_1 H_{1x} H_{0x} + m_2 H_{1z} H_{0z} - pE_{1n} E_{0n}] \\ B_1 &= \frac{j\gamma_1}{ab}[-m_1 H_{1x} H_{0x} + m_2 H_{1z} H_{0z} - pE_{1n} E_{0n}] \end{aligned}\right\} \quad . \quad (7)$$

where the subscripts 0 and 1 refer to the exciting and excited guide respectively.

For infinite directivity, $A_1 = 0$, so that with the hole in the centre of the broad face, $m_2 H_{1z} H_{0z}$ is zero and

$$m_1 H_{1x} H_{0x} = pE_{1n} E_{0n}$$

This leads to the solution, for the special case in which the guides are parallel:

$$\frac{k^2}{\gamma_0 \gamma_1} = 2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

When the hole is slightly off-centre, the term $m_2 H_{1z} H_{0z}$ is non-zero and must be included. Performing the substitutions for H and m outlined in Section 4 of Surdin's paper, we obtain

$$A_1 = \frac{j\gamma_1 p}{ab}\left[2\sin^2\left(\frac{\pi x}{a}\right) - \frac{2\alpha_0\alpha_1}{\gamma_0\gamma_1}\cos^2\left(\frac{\pi x}{a}\right) - \frac{k^2}{\gamma_0\gamma_1}\sin^2\left(\frac{\pi x}{a}\right)\right] \quad . \quad . \quad . \quad . \quad (9)$$

The hole is not exactly in the centre of the broad face and so the angle $\pi x/a$ is no longer $\pi/2$ but $(\pi/2)(1 + \delta)$, where δ is the incremental displacement. Substituting this angle in eqn. (9),

$$A_1 = \frac{j\gamma_1 p}{ab}\left[2\cos^2\left(\frac{\delta\pi}{2}\right) + \frac{2\alpha_0\alpha_1}{\gamma_0\gamma_1}\sin^2\left(\frac{\delta\pi}{2}\right) - \frac{k^2}{\gamma_0\gamma_1}\cos^2\left(\frac{\delta\pi}{2}\right)\right] \quad . \quad . \quad . \quad . \quad (10)$$

so that, for $A_1 = 0$,

$$2 + \frac{2\alpha_0\alpha_1}{\gamma_0\gamma_1}\tan^2\left(\frac{\delta\pi}{2}\right) = \frac{k^2}{\gamma_0\gamma_1} \quad . \quad . \quad . \quad (11)$$

which may be compared with eqn. (8) in which the hole is exactly in the centre. Since the directivity, D, is given by

$$D = \frac{B_1}{A_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

and

$$B_1 \simeq 2 + \frac{k^2}{\gamma_0\gamma_1} = 4 \quad . \quad . \quad . \quad . \quad . \quad (13)$$

the modified directivity, as a result of the hole being off-centre, is then

$$D_{db} = 20 \log_{10}\left[\frac{2\lambda_c^2}{\lambda_g^2 \tan^2\left(\frac{\delta\pi}{2}\right)}\right] \quad . \quad . \quad . \quad (14)$$

The directivity is only slightly affected by moderate errors in hole positioning. A hole 0·04in off-centre in a 2in guide reduces the directivity to 52dB. A stricter tolerance would probably be needed for purely mechanical reasons.

(12.2) Bandwidth of Multipost Coupler

The bandwidth of the multi-post coupler is of interest when any given application is considered. An expression derived by Tomiyasu and Cohn[7] giving the power-transfer bandwidth for total transfer is

$$\text{Bandwidth } \% \simeq \frac{200\sqrt{k}}{5\pi g h} \quad . \quad . \quad . \quad . \quad (15)$$

where k = Tolerable percentage decrease in power transfer.
$g = \lambda_{g2}/\lambda$
$h = \lambda_{g1}/\lambda$
and λ_{g1} and λ_{g2} are the guide wavelengths of the two modes.

For power transfers much less than the maximum, a more useful approximation can be derived in the following way. From eqn. (1),

$$P_3/P_1 = \sin^2 \phi \quad . \quad . \quad . \quad . \quad . \quad (16)$$

where $\phi = (\pi l_c/\lambda\lambda_c)[\sqrt{(\lambda_c^2 - \eta^2\lambda^2)} - \sqrt{(\lambda_c^2 - \lambda^2)}]$

The term η is approximately independent of wavelength. Expanding the expression about λ we have

$$\frac{\delta P_3}{\delta P_1} = \delta\lambda \frac{\pi l_c}{\lambda^3}(\lambda_{g2} - \lambda_{g1}) \sin 2\phi$$

Since, to a first order, positive and negative $\delta\lambda$ merely change the sign of δP_3,

$$\text{Bandwidth} = \frac{2\delta\lambda}{\lambda} \simeq \frac{2\lambda^2(\delta P_3/P_1)}{\pi l_c(\lambda_{g2} - \lambda_{g1}) \sin 2\phi} \quad . \quad . \quad (17)$$

Eqn. (17) is useful for estimating the bandwidth of 3 dB couplers or power transfers less than this value.

621.317.34: 621.372.85

The Institution of Electrical Engineers
Paper No. 2796 E
May 1959
©

A PERTURBATION METHOD FOR CIRCULAR WAVEGUIDES CONTAINING FERRITES

By P. J. B. CLARRICOATS, B.Sc.(Eng.), Ph.D., Graduate.

(*The paper was first received* 16*th August, and in revised form* 10*th October*, 1958.)

SUMMARY

The paper describes a perturbation method which enables the propagation coefficient of a circular waveguide, containing a longitudinally magnetized ferrite rod, to be determined over a wider range of parameters than hitherto. The method relies on the fact that, when a mode, propagating in a circular waveguide containing a longitudinally magnetized ferrite rod, is circularly polarized on the waveguide axis, the field throughout the whole of the ferrite is nearly circularly polarized. Theoretical and experimental results for the negative circularly polarized dominant mode are favourably compared in the case of a Ferroxcube B2 ferrite.

LIST OF SYMBOLS

E_t = Transverse electric field vector.
$F(x) = x\frac{J_1'(x)}{J_1(x)}$.
H_t = Transverse magnetic field vector.
l, l' = Actual and effective length of ferrite.
n = Azimuthal wavenumber.
r_1, r_0 = Rod and waveguide radius.
$\alpha, \bar{\alpha}$ = Attenuation and normalized attenuation coefficients.
$\beta, \bar{\beta}, \bar{\beta}'$ = Phase-change coefficient and normalized phase-change coefficients.
γ = Propagation coefficient.
$\epsilon_0, \epsilon, \epsilon_1$ = Permittivity of free space, ferrite or dielectric rod, and surrounding medium.
$\epsilon = \epsilon/\epsilon_0$.
κ = Off-diagonal component of permeability tensor.
$\bar{\kappa} = \kappa/\mu_0$.
$\lambda_0, \lambda, \lambda_e$ = Wavelength in free space, waveguide containing rod, and empty waveguide.
λ_c = Cut-off wavelength.
μ_0, μ, μ_1 = Permeability of free space, ferrite (diagonal component), and surrounding medium.
$\mu_t = \mu - n\kappa$ = Transverse permeability for circular polarization.
μ_z = Longitudinal permeability of ferrite.
$\bar{\mu} = \mu/\mu_0$.
ω = Angular frequency.

(1) INTRODUCTION

Although circular waveguides containing magnetized ferrite rods have been used in microwave components since 1952, their propagation behaviour has not been determined theoretically for other than a restricted range of parameters.* Because of the complexity of the propagation equation, several authors have used perturbation methods to determine the propagation coefficients. However, those described so far are only valid for thin rods or weak magnetization, neither condition being generally satisfied in microwave components. This paper describes a perturbation method which is certainly valid in the case of a negative circularly polarized wave, for rods of arbitrary radius with magnetization up to saturation. Although the validity of the method has not been tested in the case of a positive circularly polarized wave, it is expected to apply up to saturation for ferrites such as Ferramic R1, provided that the frequency exceeds about 8 Gc/s. However, experiments described in the paper show that it is the behaviour of the negative circularly polarized wave that is most important in microwave components employing longitudinally magnetized ferrites.

In Section 2, existing methods for determining the propagation coefficients are briefly reviewed and their limitations stated. Section 3 describes the present perturbation method, while in Section 4 experimental and theoretical results are compared. The experimental results are for nickel-zinc ferrite and are taken from a detailed investigation made by the author. Section 5 summarizes the perturbation method.

* After the paper was prepared, J. E. Tompkins kindly discussed with the author the initial results of a computational programme for determining the phase-change coefficients of a circular waveguide containing a Ferramic R1 ferrite rod. These results appear to confirm the results described here and in particular justify the main assumption underlying the perturbation method.

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is a communication from the Staff of the Applied Electronics Laboratories of The General Electrical Company, Limited, Stanmore, England.
Dr. Clarricoats is now in the Department of Light Electrical Engineering at the Queen's University of Belfast.

(2) EXISTING METHODS FOR DETERMINING THE PROPAGATION COEFFICIENTS

The propagation equation for a circular waveguide containing a magnetized ferrite rod of arbitrary size has been given by Kales.[1] Because of the complicated transcendental form of the equation, solutions can, in general, be obtained only by numerical methods. Rizzi[2] has thus evaluated the propagation coefficients, for negative and positive circularly polarized waves, of a waveguide containing a Ferramic R1 ferrite rod. Owing to the labour involved, Rizzi gives results for only two rod diameters and with parameters appropriate to magnetic saturation.

Van Trier[3] and Suhl and Walker[4] have derived an expression for the change in the propagation coefficient of a circular waveguide, due to the introduction of a thin magnetized ferrite rod. Table 1 shows the ratio r_1/r_0, of the rod to the waveguide radius, which produces a 10% change in propagation coefficient. The values refer to nickel-zinc (Ferroxcube B2) and magnesium-manganese (Ferramic R1) ferrites at magnetic saturation. The results in Table 1 are based on measurements of the ferrite permeability and permittivity made by MacBean[5] and LeCraw and Spencer.[6] Fig. 1 indicates that when r_1/r_0 exceeds the value at which the thin-rod perturbation $\delta\beta/\beta$ is 10%, the difference between the two propagation curves increases rapidly with r_1/r_0. If the results of Section 3 (Fig. 2) are anticipated, and curve (*a*) of Fig. 1 is assumed to be nearly correct, it may be concluded that the thin-rod perturbation rapidly becomes invalid when it exceeds 10%. Since most microwave components employ ferrite rods with $r_1/r_0 > 0{\cdot}25$, it is evident that the thin-rod perturbation is of slight practical value.

Suhl and Walker[7] have also considered the perturbation of the

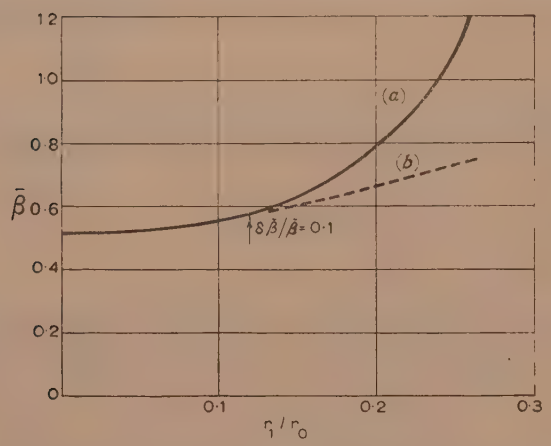


Fig. 1.—Calculated normalized phase-change coefficients, $\bar{\beta} = \lambda_0/\lambda$, for negative circular polarization, as a function of the ratio r_1/r_0.

Ferrite and waveguide parameters: $\bar{\mu} = 0{\cdot}7$, $\bar{\kappa} = 0{\cdot}9$, $\bar{\mu}_z = 0{\cdot}9$, $\bar{\epsilon} = 11{\cdot}4$ Ferroxcube B2) $2r_0/\lambda_0 = 0{\cdot}685$ (0·875 in diameter waveguide at 9·25 Gc/s).

(*a*) $\bar{\beta}$ calculated from Section 3 perturbation.
(*b*) $\bar{\beta}$ calculated from thin rod perturbation.

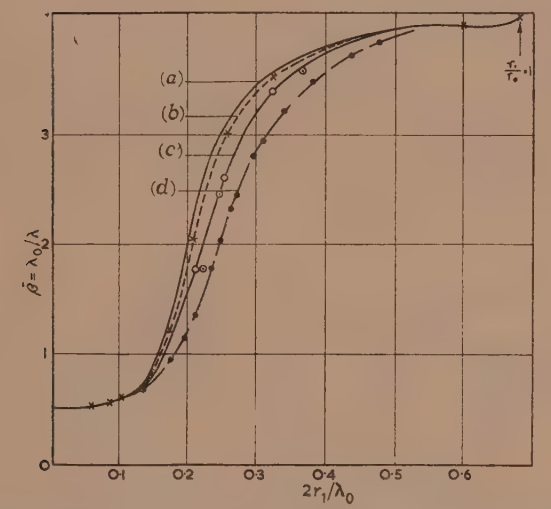


Fig. 2.—Theoretical and experimental propagation curves for H_{11}-type mode in a waveguide partially filled with magnetized ferrite.

$\bar{\mu}_t = 1{\cdot}6$, $\bar{\kappa} = 0{\cdot}9$, $\bar{\epsilon} = 11{\cdot}4$, $2r_0/\lambda_0 = 0{\cdot}685$, $\bar{\mu}_z = 0{\cdot}9$.
(*a*) ———— Isotropic permeability, $\bar{\mu}_t = \bar{\mu}_z = \bar{\mu} + \bar{\kappa} = 1{\cdot}6$.
(*b*) —×—×— Scalar permeability, $\bar{\mu}_t = 1{\cdot}6$, $\bar{\mu}_z = 0{\cdot}9$.
(*c*) —○—○— Perturbation of scalar permeability, $\bar{\mu} = 0{\cdot}7$, $\bar{\kappa} = 0{\cdot}9$, $\bar{\mu}_z = 0{\cdot}9$.
(*d*) -●-●-●- Experimental results for sample 0·78 in long.
—⊙—⊙— Experimental results allowing for end corrections.

propagation coefficient, $\delta\beta$, produced by the slight magnetization of a ferrite rod contained in an otherwise empty waveguide. Their final expression for $\delta\beta$ is given as

$$\delta\beta = (\mu - \mu_0)P - n\kappa Q \quad . \quad . \quad . \quad . \quad (1)$$

where P and Q are defined by

$$P = C\int_{rod} \boldsymbol{H}_t \,.\, \boldsymbol{H}_t^* ds$$

$$Q = -\frac{jC}{n}\int_{rod} (\boldsymbol{i}_z \times \boldsymbol{H}_t) \,.\, \boldsymbol{H}_t^* ds$$

Table 1

VALUES OF THE RATIO r_1/r_0, TO PRODUCE A 10% PERTURBATION OF THE PROPAGATION COEFFICIENT WHEN $2r_0/\lambda_0 = 0{\cdot}685$

Ferrite	Negative circular polarization†	Positive circular polarization†
Ferroxcube B2 ..	0·12	0·17
Ferramic R1 ..	0·11	0·13

† In accordance with the convention used in describing ferrite behaviour, a positive circularly polarized wave has a field pattern which has a right-handed rotation with respect to the direction of the static magnetic field. This convention differs from that used in reciprocal systems, e.g. microwave antennae, where polarization sense is defined with respect to the direction of propagation.

where $\boldsymbol{H}_t$ is the unperturbed transverse magnetic field and $\boldsymbol{H}_t^*$ is its complex conjugate, C is a function depending on r_1/r_0, r_0/λ_0 and ϵ, and $\boldsymbol{i}_z$ is a unit vector in the z-direction. In this paper, it is assumed that the ferrite rod is magnetized in the positive z-direction along which the wave propagates. All cylindrical field components are assumed to have a dependence $e^{j(\omega t - \beta z - n\theta)}$, the whole field pattern thus rotating about the axis at a rate $\theta/t = \omega/n$. A wave is positively or negatively circularly polarized on the waveguide axis, according as n is $+1$ or -1.

It is assumed in eqn. (1) that the unmagnetized ferrite has the free-space permeability μ_0; this is only justified when frequencies above about 18 Gc/s are being considered. At 9·25 Gc/s the relative permeability, $\mu/\mu_0 = \bar{\mu}$, of unmagnetized Ferroxcube B2 ferrite is 0·2, while that of Ferramic R1 ferrite is 0·75. The above authors show P and Q as functions of r_1/r_0 for a ferrite with $\bar{\epsilon} = 10$ and $2r_0/\lambda_0 = 0{\cdot}8$. From these results the present author has determined that, for a 10% perturbation, $(\mu \pm \kappa - \mu_0)/\mu_0 = |\bar{\mu} \pm \bar{\kappa} - 1|$ must lie between 0·1 and 0·15 for r_1/r_0 between 0·2 and 0·9, and for $r_1/r_0 = 1{\cdot}0$, $\bar{\mu} \pm \bar{\kappa} - 1$ must not exceed 0·3. Since in the unmagnetized and magnetically saturated states $|\bar{\mu} \pm \bar{\kappa} - 1|$ has values[5,6] of 0·8 and 0·6 respectively for Ferroxcube B2 ferrite, and 0·3 and 0·45 for Ferramic R1 ferrite, it would appear that this perturbation has a restricted practical value.* It is not possible in this case to make a comparison with curve (*a*) of Fig. 1, since, for slight magnetization values, Ferroxcube B2 has a large magnetic loss tangent and the above perturbation method cannot readily be applied.

(3) PERTURBATION OF WAVEGUIDE CONTAINING A ROD WITH SCALAR TRANSVERSE PERMEABILITY

(3.1) Description of Perturbation Method

It is well known that, when magnetized in the direction of propagation, a ferrite presents a scalar permeability $\mu_t = \mu \pm \kappa$ to a circularly polarized transverse r.f. magnetic field. The plus sign is associated with negative circular polarization, and the minus sign with positive circular polarization. The complication in the propagation equation of a waveguide containing a magnetized ferrite rod arises partly because the transverse magnetic field within the ferrite is not circularly polarized over the whole cross-section when circularly polarized on the axis. However, the perturbation to be described here was inspired by the knowledge that the effect of the deviation of the field from circular polarization is in many cases slight. For example, the difference between the quantities P and Q in eqn. (1) is a measure of the extent to which the transverse magnetic field within a dielectric rod deviates from being circularly polarized. When $\bar{\epsilon} = 10$, the difference does not exceed 15% in the range $0 < r_1/r_0 < 1$.

* Both Ferroxcube B2 and Ferramic R1 ferrites have relative permittivities greater than 10; therefore the results of Suhl and Walker cannot be used directly in the case of these ferrites.

The perturbation considered is that which arises when the permeability tensor changes from the diagonal form

$$\begin{bmatrix} \mu - n\kappa & 0 & 0 \\ 0 & \mu - n\kappa & 0 \\ 0 & 0 & \mu_z \end{bmatrix} \text{ to the non-diagonal form } \begin{bmatrix} \mu & -j\kappa & 0 \\ j\kappa & \mu & 0 \\ 0 & 0 & \mu_z \end{bmatrix}$$

A first stage in the application of this perturbation method is to determine the propagation coefficients of a waveguide containing a rod with different transverse and longitudinal permeabilities. Fortunately the propagation equation is only slightly more complicated than that of a waveguide containing an isotropic dielectric rod.* A second stage is to obtain an expression, similar to eqn. (1), for the change in propagation coefficient caused by a change in the form of the tensor permeability.

(3.2) The Propagation Equation for a Waveguide containing a rod with Different Transverse and Longitudinal Permeabilities

It can be readily shown that the equation for the propagation coefficient, γ, of a waveguide containing a rod with transverse permeability, μ_t, and longitudinal permeability, μ_z, is given by

$$-\frac{n^2\gamma^2}{\omega^2}\frac{[(K_1r_1)^2-(K'r_1)^2]^2}{[(K_1r_1)^2(K'r_1)^2]^2} = \left[\frac{\mu_t F(K''r_1)}{(K'r_1)^2} - \frac{\mu_1 R(K_1r_1)}{(K_1r_1)^2}\right] \times \left[\frac{\epsilon F(K'r_1)}{(K'r_1)^2} - \frac{\epsilon_1 S(K_1r_1)}{(K_1r_1)^2}\right] \quad . \quad (2)$$

where μ_1 and ϵ_1 are respectively the permeability and permittivity of the surrounding medium and

$$(K')^2 = \omega^2\epsilon\mu_t + \gamma^2,\ (K'')^2 = \frac{\mu_z}{\mu_t}(K')^2,\ K_1^2 = \omega^2\epsilon_1\mu_1 + \gamma^2$$

$$F(x) = x\frac{J_1'(x)}{J_1(x)}$$

$$R(K_1r_1) = (K_1r_1)\frac{[J_1'(K_1r_1)Y_1'(K_1r_0) - J_1'(K_1r_0)Y_1'(K_1r_1)]}{[J_1(K_1r_1)Y_1'(K_1r_0) - J_1'(K_1r_0)Y_1(K_1r_1)]}$$

$$S(K_1r_1) = (K_1r_1)\frac{[J_1'(K_1r_1)Y_1(K_1r_0) - J_1(K_1r_0)Y_1'(K_1r_1)]}{[J_1(K_1r_1)Y_1(K_1r_0) - J_1(K_1r_0)Y_1(K_1r_1)]}$$

Eqn. (2) has been solved,* as a function of r_1/r_0, for the following rod and waveguide parameters: $\bar{\mu}_t = 1{\cdot}6$, $\bar{\mu}_z = 0{\cdot}9$, $\bar{\epsilon} = 11{\cdot}4$, $\bar{\mu}_1 = \bar{\epsilon}_1 = 1$, $2r_0/\lambda_0 = 0{\cdot}685$. Both the rod and surrounding medium were assumed lossless. Considerable labour can be saved if the propagation coefficient is known as a function of r_1/r_0 when $\bar{\mu}_t = \bar{\mu}_z = 1{\cdot}6$.* Curves showing the normalized phase-change coefficient $\bar{\beta} = \lambda_0/\lambda$ as a function of r_1/r_0 for the H_{11}-type mode are contained in Fig. 2 [Curves (*a*) and (*b*)]. The ferrite parameters, chosen above, correspond to those of a magnetically saturated Ferroxcube B2 ferrite placed in a negative circularly polarized field. This type of ferrite was selected for the calculations in order that they might be compared with the results of an experimental study. At the time this study was made, only Ferroxcube B2 ferrite was available in suitable sizes. It was decided to make the comparison with the rod just saturated, to avoid complications, arising in the interpretation of experiments, due to multi-mode interference.[8]

In the special case of a completely filled waveguide, exact expressions for the phase-change and attenuation coefficients have been given previously.[8] These are

$$(\bar{\beta}, \bar{\alpha}) = \sqrt{\left\{\pm\frac{1}{2}\left[\bar{\epsilon}\bar{\mu}_t' - \left(\frac{\lambda_0}{\lambda_c}\right)^2\left(\frac{\mu_t'\mu_z' + \mu_t''\mu_z''}{|\mu_z|^2}\right)\right] + \frac{1}{2}\frac{|\mu_t|}{|\mu_z|}\sqrt{\left[\bar{\epsilon}^2|\bar{\mu}_z|^2 - 2\left(\frac{\lambda_0}{\lambda_c}\right)^2\bar{\epsilon}\bar{\mu}_z' + \left(\frac{\lambda_0}{\lambda_c}\right)^4\right]}\right\}} \quad . \quad (3)$$

where it is to be understood that the plus sign refers to $\bar{\beta}$ and the minus to $\bar{\alpha}$. In the case of H-modes, $\lambda_c = 2\pi r_0/u_{nm}$ where u_{nm} is the mth root of the Bessel function $J_n(x)$. The normalized attenuation coefficient $\bar{\alpha}$ is defined by $\bar{\alpha} = \lambda_0\alpha/2\pi$. α is the attenuation coefficient appearing in the relation $\gamma = \alpha + j\beta$. For simplicity, dielectric loss is assumed zero in the derivation of eqn. (3), although no difficulty arises with its introduction. If magnetic loss is also negligible, a particularly simple expression for the phase-change coefficient results:

$$\bar{\beta} = \sqrt{\left[\bar{\epsilon}\bar{\mu}_t - \left(\frac{\lambda_0}{\lambda_c}\right)^2\frac{\mu_t}{\mu_z}\right]} \quad . \quad . \quad . \quad . \quad (4)$$

Eqn. (4) has been used to calculate the phase-change coefficient of the dominant H-type mode in a waveguide filled with a saturated Ferramic R1 ferrite. The results of this calculation are compared in Table 2 with those obtained by Rizzi, who solved the exact transcendental equation by numerical methods. It must be noted that the parameters he employed do not correspond precisely with the measured values for R1 ferrite. For comparison purposes, Rizzi's values were inserted in eqn. (4).

Table 2

COMPARISON OF THEORETICAL VALUES OF NORMALIZED PHASE-CHANGE COEFFICIENT FOR NEGATIVE AND POSITIVE CIRCULAR POLARIZATION

Normalized phase-change coefficient			
Exact solution		Eqn. (4)	
Sense of polarization			
Negative	Positive	Negative	Positive
3·79	1·94	3·86	1·78

Parameters: $\bar{\mu} = 0{\cdot}95$, $\bar{\kappa} = 0{\cdot}62$, $\bar{\mu}_z = 1{\cdot}0$, $\bar{\epsilon} = 10$, $2r_0/\lambda_0 = 0{\cdot}318$.

For negative circular polarization the agreement between values of $\bar{\beta}$ is 1·3%, and for positive, 9%. In view of the considerable difference in the labour of the two calculations, there now seems slight justification for using the exact equation, other than when higher-order modes are being considered or when internal static magnetic fields much greater than the saturation value are employed. The analysis made by Suhl and Walker suggests that in these cases the r.f. fields will not be circularly polarized over most of the cross-section. It is also consistent with their analysis that the agreement in Table 2 is best for negative circular polarization, since they noted that it was for this sense of polarization that the ferrite mode most resembled the H_{11}-mode in isotropic waveguide. For the same reason, eqn. (4) becomes inaccurate for positive circular polarization when $\bar{r}_0 = (2r_0/\lambda_0)\pi\sqrt{(\bar{\epsilon}\bar{\mu}_z)} > 4$; in the above case $\bar{r}_0 \simeq 3$.

(3.3) Perturbation from Scalar to Tensor Permeability

The following is an expression[10] for the perturbation, $\delta\beta$, due to the change in the permeability of the rod region s_1 from

* A detailed discussion of the solution of both types of equation will be presented in a later paper.

the diagonal form involving $\mu_t = \mu - n\kappa$ to the tensor form involving μ and κ:

$$\delta\beta = \frac{\omega}{2\int_s (E_t \times H_t^*) \cdot i_z ds} \times \int_{s_1} [(\mu - \mu_t) H_{to} \cdot H_t^* + j\kappa (i_z \times H_{to}) \cdot H_t^*] ds \quad . \quad (5)$$

where the suffix o denotes the perturbed field and $\mu - \mu_t = n\kappa$. Provided that the perturbation is sufficiently small, the perturbed fields may be made equal to those originally within the rod. Replacing the integral in the denominator by an expression involving the slope of the unperturbed propagation coefficient curve $d\bar{\beta}/dr_1$, and evaluating the integrals in the numerator, $\delta\bar{\beta}$ is obtained as

$$\delta\bar{\beta} = \frac{\bar{\kappa} r_1 \dfrac{d\bar{\beta}}{dr_1} \{ n[\bar{\Lambda}^2 \bar{\beta}^2 I_1 + \bar{\epsilon}^2 I_2 + 2n\bar{\epsilon}\bar{\Lambda}\bar{\beta} I_3]}{(\bar{\epsilon} - 1)\{ J_1^2(y')(\bar{\epsilon}\bar{\mu}_t - \bar{\beta}^2)^2 (y')^2 + \bar{\epsilon}[\bar{\beta} J_1(y') + n\bar{\mu}_t \bar{\Lambda} J_1(y'')]^2 + [n\bar{\beta} J_1(y') + \bar{\mu}_t \bar{\Lambda} y'' J_1'(y'')]^2 \}}$$

$$\frac{- [n\bar{\Lambda}^2 \bar{\beta}^2 J_1(y'') + n\bar{\epsilon}^2 J_1^2(y') + 2\bar{\epsilon}\bar{\Lambda}\bar{\beta} I_4] \}}{+ (\bar{\mu}_z - 1)\bar{\Lambda}^2 J_1^2(y'')(\bar{\epsilon}\bar{\mu}_t - \bar{\beta}^2)^2 (y')^2 + (\bar{\mu}_t - 1)\{ \mu_t [\bar{\beta}\bar{\Lambda} y'' J_1'(y'') + n\bar{\epsilon} J_1(y')]^2 + [n\bar{\beta}\bar{\Lambda} J_1(y'') + \bar{\epsilon} J_1(y')]^2 \}} \quad . \quad (6)$$

where $\bar{\beta}$ is a solution of eqn. (2) with $\bar{\mu}_1 = \bar{\epsilon}_1 = 1$,

$$y = Kr_1, \; y' = K'r_1, \; y'' = K''r_1$$

$$\bar{\Lambda} = \frac{J_1(y')\left[\dfrac{S(y)}{y^2} - \dfrac{\bar{\epsilon} F(y')}{(y')^2}\right]}{J_1(y'')n\bar{\beta}\left[\dfrac{1}{(y')^2} - \dfrac{1}{y^2}\right]}$$

$$I_1 = J_1^2(y')\{\tfrac{1}{2}[(y')^2 - 1] + \tfrac{1}{2}F^2(y') + F(y')\}$$

I_2 has the same form as I_1 with y'' replacing y'

$$I_3 = J_1(y')J_1(y'')$$

$$I_4 = \frac{I_3[F(y')(y'')^2 - F(y'')(y')^2]}{(y'')^2 - (y')^2}$$

In spite of the formidable appearance of eqn. (6), relatively few functions are involved and the calculation of $\delta\bar{\beta}$ is straightforward. Fig. 2, curve (*c*), shows results for negative circular polarization obtained by applying the perturbation to a rod with the same parameters as in curve (*b*), with the additional specification $\bar{\mu} = 0{\cdot}7, \bar{\kappa} = 0{\cdot}9$. The results in curve (*b*) were for scalar transverse permeability $\bar{\mu}_t = \bar{\mu} + \bar{\kappa} = 1{\cdot}6$. These parameters are taken from measurements made on Ferroxcube B2 ferrite by MacBean.[5] The largest percentage difference between curves (*b*) and (*c*) is 15%, and over a wide range of r_1/λ_0 it is less than 10%. A 15% perturbation might be considered unreasonably large if the object were a precise determination of ferrite parameters. However, since the corresponding error in $\bar{\beta}$ probably does not exceed a few per cent, the method is considered quite satisfactory for the design of microwave components. The agreement between experiment and theory, discussed in Section 4, supports this claim. Furthermore, since the real parts of the permeability and permittivity of Ferroxcube B2 ferrite are typical of many other ferrites used in microwave components, the above method appears quite suitable for determining their corresponding propagation coefficients. Also, since most of these ferrites have loss tangents of at least an order less than those of Ferroxcube B2, the neglect of loss in the perturbation calculation should cause a smaller error. Since the above perturbation is reasonable at magnetic saturation and correctly zero for the unmagnetized ferrite (for which $\bar{\kappa} = 0$), it is to be expected that it will serve for all magnetizations up to saturation.

The present perturbation is now contrasted with that produced by slight magnetization. The perturbation expression was given in eqn. (1) by $\delta\beta = (\mu - \mu_0)P + \kappa Q$, where P and Q were approximately equal quantities. Eqn. (6), on the other hand, is of the form $\delta\beta = \kappa(P' - Q')$, where P' and Q' are also approximately equal and of similar magnitudes to P and Q. Clearly, at a given rod radius, for $\delta\beta$ in eqns. (1) and (6) to be the same, κ is larger in the latter case, i.e. the present perturbation is valid for larger magnetization. As in the prevous perturbation, the difference between P' and Q' is a direct measure of the deviation from circular polarization of the unperturbed transverse magnetic field.

(4) DISCUSSION ON THEORETICAL AND EXPERIMENTAL RESULTS

Fig. 2, curve (*d*), shows an effective normalized propagation coefficient, $\bar{\beta}'$, determined from the results of an experimental investigation. The phase shift, attenuation and reflection caused by the introduction of a Ferroxcube B2 ferrite rod into a circular waveguide were measured using a microwave bridge.[8,10] The range of applied field extended from 0 to 3 500 gauss and both senses of circular polarization were employed. The results in Fig. 2 were obtained using one sample, the rod diameter being reduced from 0·625 in by grinding. A disc, cut from the same rod, was employed in the measurement of μ and κ.

The effective phase-change coefficient, $\bar{\beta}'$ is related to the measured differential phase shift, ϕ_s (relative to the empty waveguide), expressed as a multiple of 2π and the rod length l, by the equation

$$\bar{\beta}' = \frac{\lambda_0}{\lambda_e} + \frac{\lambda_0}{l}\phi_s \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where λ_e is the wavelength in empty waveguide.

Eqn. (7) cannot be expected to give the same phase-change coefficient as obtained using the perturbation method, since that value refers to an infinitely long rod, whereas the above is for a rod of finite length. In order to take account of the rod ends completely one would have to observe ϕ_s as a function of rod length for different r_1/r_0 ratios throughout the range. However, it was considered adequate to make measurements at three values of r_1/r_0 as shown in Fig. 3.

It can be seen from Fig. 3 that the phase shift depends almost linearly upon the rod length. This suggests that end effects are practically independent of length although dependent on diameter. On denoting the intercept with the rod length axis by l', the phase-change coefficient $\bar{\beta}$ is obtained as

$$\bar{\beta} = \frac{\lambda_0}{\lambda_e} + \frac{\lambda_0 \phi_s}{l - l'} \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Values of $\bar{\beta}$ obtained using eqn. (8) have been inserted in Fig. 2. It can be seen that for two values of r_1/r_0 the agreement with curve (*c*) is better than 2%, while for one the agreement is 10%.

It is not possible to assess the overall experimental accuracy

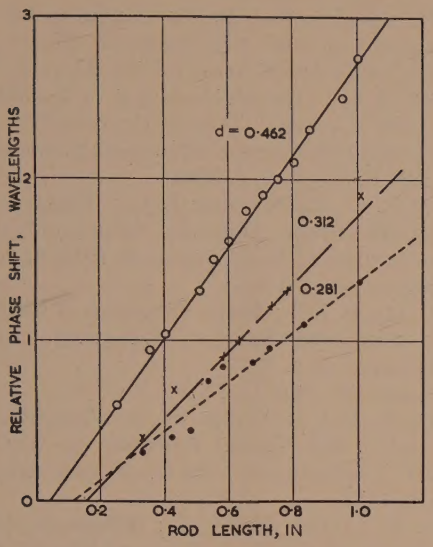


Fig. 3.—Relative phase shift as a function of rod length.

The B2 rods are supported by foam. Parameter: rod diameter d in inches. Waveguide diameter = 0·875 in, frequency = 9·25 Gc/s.

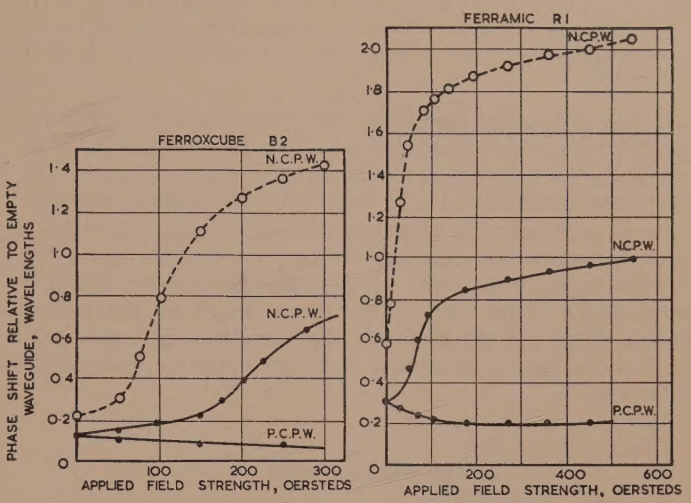


Fig. 4.—Phase shift of positive and negative circularly polarized waves for foam-supported Ferroxcube B2 and Ferramic R1 rods.

—— Rod length 1 in, ---- Rod length 2 in.
Diameter of rod = 0·25 in, diameter of waveguide = 0·875 in.
Frequency = 9·25 Gc/s.
N.C.P.W. = Negative circularly polarized wave.
P.C.P.W. = Positive circularly polarized wave.

with precision since a large number of different factors are involved. The most important of these appear to be:

(*a*) Differences between the parameters of the ferrites used in the investigation of the effects of variation of diameter and length. Although the four ferrite samples were all nominally of the same composition, experiments have shown that even ferrites from the same firing batch show variations in parameters.

(*b*) Effects arising from interference between the H_{11}-type and E_{11}-type modes. Other experiments[10] indicate that these effects are slight when the ferrite is magnetically saturated.

(*c*) Errors in the measurement of the parameters μ, κ and ϵ. The estimated accuracy is quoted as 3%.[5]

Of these factors, the first is probably mainly responsible for the 10% error with the rod of diameter 0·281 in. For the other two rod diameters the agreement lies well within the experimental accuracy, which is probably not less than 5%.

It is noteworthy that, if a reduction in the accuracy of $\bar{\beta}$ can be tolerated, one can perturb about an isotropic dielectric rod with permeability $\mu + \kappa$ [Fig. 2, curve (*a*)]. Although this implies a maximum perturbation, in the above case, of 25%, the maximum error has been found not to exceed 16% assuming curve (*c*) to be correct. Also, the perturbation expression is simpler and the propagation coefficient for the isotropic rod can be obtained more easily.

In spite of the uncertainty regarding the experimental accuracy, the results in Fig. 2 appear to confirm the usefulness of the present perturbation method when applied to the negative circularly polarized mode. It remains to be shown whether the method will yield satisfactory results for the positive circularly polarized mode with Ferroxcube B2, since the scalar permeability at saturation is then negative. However, for ferrites such as Ferramic R1, $\bar{\mu}_t$ is still positive at saturation for positive circular polarization, so the method should yield results comparable with those above. A comparison awaits further experimental or theoretical results for Ferramic R1. This shortcoming is not serious, since Fig. 4 indicates that the phase change at saturation for negative circular polarization well exceeds that for positive, and thus the former would normally be used in circularly polarized wave phase-shifter applications. Also, in Faraday rotation applications, when r_1/r_0 exceeds about 0·2, rotation arises mainly from the phase shift of the negative circularly polarized component of the linearly polarized wave. Under such conditions the rotation produced by a ferrite of length l is approximately $\frac{1}{2}(\bar{\beta}_- l)$, where $\bar{\beta}_-$ is the phase-change coefficient of the negative circularly polarized mode.

Certain ferrites when magnetized to saturation have $\mu - \kappa \simeq 0$, and for these the phase-change coefficient of the positive circularly polarized mode, $\bar{\beta}_+$, can be quite readily determined without the need for a perturbation method. This follows since a number of authors[10, 11, 12] have shown that a waveguide containing a ferrite rod with $\mu_t \simeq 0$ behaves similarly to a waveguide containing a metallic inner conductor (i.e. a coaxial line). The phase-change coefficient is then obtainable from the propagation equation for a coaxial line.[13]

(5) CONCLUSIONS

In conclusion, the stages leading to the final propagation curve in Fig. 2 [curve (*c*)], are summarized:

(*a*) The transcendental equation (eqn. 2) for the propagation coefficient of a waveguide containing a rod with parameters $\mu_t = \mu - n\kappa$, μ_z, and ϵ, is solved by conventional numerical methods. As a result curve (*b*) of Fig. 2 is obtained.

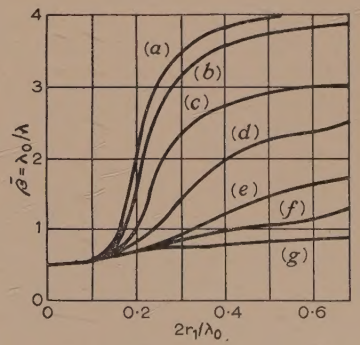


Fig. 5.—$\bar{\beta}$ as a function of $2r_1/\lambda_0$ for H_{11}-type mode in a waveguide containing a dielectric rod.

Parameters: $2r_0/\lambda_0 = 0{\cdot}685$.
(*a*) $\bar{\epsilon}\bar{\mu} = 18{\cdot}4$. (*b*) $\bar{\epsilon}\bar{\mu} = 16$. (*c*) $\bar{\epsilon}\bar{\mu} = 10$. (*d*) $\bar{\epsilon}\bar{\mu} = 7$. (*e*) $\bar{\epsilon}\bar{\mu} = 4$. (*f*) $\bar{\epsilon}\bar{\mu} = 2{\cdot}5$. (*g*) is for a coaxial line supporting the H_{11} mode. Diameter of filled waveguide = 0·875 in.

(*b*) The slope $d\bar{\beta}/dr_1$ of curve (*b*) is determined as a function of r_1.

(*c*) The perturbation $\delta\bar{\beta}$ for a change from scalar to tensor permeability is evaluated using eqn. (6) and curve (*c*) is thus obtained.

It was previously mentioned that the labour in obtaining curve (*b*) is considerably reduced if the propagation curve is available for a rod with isotropic permeability $\mu - n\kappa$ [curve (*a*)]. This curve does not need to be highly accurate, it merely provides a useful trial value of $\bar{\beta}$ in the solution of eqn. (2). Fig. 5 shows a number of such approximate propagation curves for different values of the parameter $\bar{\mu}\bar{\epsilon}$, where $\bar{\mu}$ is here the isotropic permeability of the rod. It has been shown by the author elsewhere that it is the product, rather than the individual values of $\bar{\mu}$ and $\bar{\epsilon}$, that is important in determining the propagation behaviour.

(6) REFERENCES

(1) KALES, M. L.: 'Modes in Waveguides Containing Ferrites', *Journal of Applied Physics*, 1953, **24**, p. 604.

(2) RIZZI, P. A.: 'The Electromagnetic Field in Composite Dielectric and Ferrite Waveguide Structures', D.Eng. Thesis, Yale University, 1953.

(3) TRIER, A. A. VAN: 'Guided Electromagnetic Waves in Anisotropic Media', *Applied Scientific Research*, 1953, **3**, p. 385.

(4) SUHL, H., and WALKER, L. R.: 'Topics in Guided Wave Propagation through Gyromagnetic Media', Part III, *Bell System Technical Journal*, 1954, **33**, p. 1142.

(5) MACBEAN, I. G.: 'The Measurement of Complex Permittivity and Complex Tensor Permeability of Ferrite Materials at Microwave Frequencies', *Proceedings I.E.E.*, 1957 (**104** B, Suppl. 6, p. 296).

(6) LE CRAW, R. C., and SPENCER, E. G.: 'Tensor Permeability of Ferrites below Magnetic Saturation', *Convention Record Institute of Radio Engineers*, 1956, **4**, Part 5, p. 66.

(7) See Reference (4), p. 1145.

(8) CLARRICOATS, P. J. B.: 'Some Properties of Circular Waveguides containing Ferrites', *Proceedings I.E.E.*, 1957 (**104** B, Suppl. 6, p. 286).

(9) See Reference (4), Part I, p. 579.

(10) CLARRICOATS, P. J. B.: 'Properties of Waveguides Containing Ferrites with Special Reference to Waveguides of Circular Cross-Section', Ph.D. Thesis, University of London, 1958.

(11) DUNCAN, B. J., and SWERN, L.: 'Effects of Zero Ferrite Permeability on Circularly Polarized Waves', *Proceedings of the Institution of Radio Engineers*, 1957, p. 647.

(12) TOMPKINS, J. E.: 'Energy Distribution in Partially Ferrite Filled Waveguides', *Journal of Applied Physics*, 1958, **29**, Part III, p. 399.

(13) MARCUVITZ, N.: 'The Waveguide Handbook' (McGraw-Hill, 1951), p. 77.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. ELECTRONIC AND COMMUNICATION ENGINEERING (INCLUDING RADIO ENGINEERING), MAY 1959

CONTENTS

	PAGE
The Acoustic Design of Talks Studios and Listening Rooms C. L. S. GILFORD, M.Sc.	245
Observations on the U.S.S.R. Earth Satellites and the Study of Radio-Wave Propagation. W. C. BAIN, M.A., B.Sc., Ph.D., and E. D. R. SHEARMAN, B.Sc.(Eng.)	259
Transistors and Associated Semiconductor Devices (Progress Review) R. G. HIBBERD, B.Sc.	264
The Application of Transistors to Line Communication Equipment. H. T. PRIOR, B.Sc., D. J. R. CHAPMAN, B.Sc., and A. A. M. WHITEHEAD, B.Sc.	279
The Application of Junction Transistors to Audio-Frequency Telephone Line Amplifiers. H. G. BASSETT, B.Sc.(Eng.), D. THOMSON and P. E. GREENAWAY, B.Sc.	290
Discussion on the above two papers	294
A Design for an Automatic Graph Plotter M. P. ATKINSON, B.Eng., W. T. BANE, B.Sc., and D. L. A. BARBER, B.Sc.	299
The Cylindrical Cross-Capacitor as a Calculable Standard A. M. THOMPSON, B.Sc.	307
A Method for the Accurate Measurement of Mutual Conductance of Thermionic Valves M. R. CHILD and D. J. SARGENT, B.Sc.	311
The Performance of Directive Aerials in Complex U.H.F. Fields J. A. SAXTON, D.Sc., Ph.D., and B. N. HARDEN, M.Sc.	315
Hemi-Isotropic Radiators for the S- or X-Band E. G. A. GOODALL, M.Sc.	318
The Design and Testing of Integrally Constructed Waveguide Assemblies G. CRAVEN and V. H. KNIGHT	321
A Perturbation Method for Circular Waveguides containing Ferrites P. J. B. CLARRICOATS, B.Sc.(Eng.), Ph.D.	335

Declaration on Fair Copying.—Within the terms of the Royal Society's Declaration on Fair Copying, to which The Institution subscribes, material may be copied from issues of the *Proceedings* (prior to 1949, the *Journal*) *which are out of print and from which reprints are not available.* The terms of the Declaration and particulars of a Photoprint Service afforded by the Science Museum Library, London, are published in the *Journal* from time to time.

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—SMITH, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 4001 R, December, 1954 (102 B, p. 1234).

The Benevolent Fund

Have YOU yet responded to the appeal for contributions to the

HOMES FUND

THE INSTITUTION'S WAR MEMORIAL

Sixteen houses have been built but there is still space for a further ten houses on 'The Chesters' Estate

The Court of Governors hope that every member will contribute to this worthy object

Contributions may be sent by post to

THE INCORPORATED BENEVOLENT FUND OF THE INSTITUTION OF ELECTRICAL ENGINEERS, SAVOY PLACE, LONDON, W.C.2

or may be handed to one of the Local Hon. Treasurers of the Fund

Local Hon. Treasurers of the Fund:

EAST MIDLAND CENTRE *R. C. Woods*
IRISH BRANCH *A. Harkin, M.E.*
MERSEY AND NORTH WALES CENTRE *D. A. Picken*
NORTH-EASTERN CENTRE . . . *J. F. Skipsey, B.Sc.*
NORTH MIDLAND CENTRE *E. C. Walton, Ph.D., B.Eng.*
SHEFFIELD SUB-CENTRE *F. Seddon*
NORTH-WESTERN CENTRE . *E. G. Taylor, B.Sc.(Eng.)*
NORTH LANCASHIRE SUB-CENTRE *G. K. Alston, B.Sc.(Eng.)*
NORTHERN IRELAND CENTRE . . *G. H. Moir, J.P.*
SCOTTISH CENTRE *R. H. Dean, B.Sc.Tech.*
NORTH SCOTLAND SUB-CENTRE . . . *P. Philip*
SOUTH MIDLAND CENTRE . *Capt. J. H. Patterson, R.A.*
RUGBY SUB-CENTRE *P. G. Ross, B.Sc.*
SOUTHERN CENTRE *G. D. Arden*
WESTERN CENTRE (BRISTOL) . . . *A. H. McQueen*
WESTERN CENTRE (CARDIFF) . . . *E. W. S. Watt*
WEST WALES (SWANSEA) SUB-CENTRE *O. J. Mayo*
SOUTH-WESTERN SUB-CENTRE . . *W. E. Johnson*

Members are asked to bring to the notice of the Court of Governors any deserving cases of which they may have knowledge

THE BENEVOLENT FUND

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* COVent Garden 1871. *Telegrams:* 'Voltampere, Phone, London.
Printed by Unwin Brothers Limited, Woking and London.